Philosophical Analysis

Edited by
Katherine Dormandy, Rafael Hüntelmann,
Christian Kanzian, Uwe Meixner,
Richard Schantz, Erwin Tegtmeier

Volume 81

Quo Vadis, Metaphysics?

Essays in Honor of Peter van Inwagen

Edited by
Mirosław Szatkowski

DE GRUYTER

ISBN 978-3-11-066249-8
e-ISBN (PDF) 978-3-11-066481-2
e-ISBN (EPUB) 978-3-11-066264-1
ISSN 2198-2066

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

© 2019 Walter de Gruyter GmbH, Berlin/Boston
Printing and binding: CPI books GmbH, Leck.

www.degruyter.com

Contents

Acknowledgements —— XI

The Declaration of the Establishment of the *International Society for Formal Ontology* —— XIII

Mirosław Szatkowski
Introduction —— 1

Part I **Peter van Inwagen – A Profile**

Mirosław Szatkowski and Peter van Inwagen
Interview with Peter van Inwagen —— 11

Elisabeth M. Bolduc
Jonathan and Peter —— 33

Part II **The Grounds and Ways of Metaphysics**

Gila Sher
Where Are You Going, Metaphysics, and How Are You Getting There? – Grounding Theory as a Case Study —— 37

Benjamin Schnieder
On the Relevance of Grounds —— 59

Uwe Meixner
Metaphysical Differences —— 83

Anna-Sofia Maurin
A Van Inwagenian Defense of Constitutionalism —— 103

Carl J. Posy
Inside the Metaphysical Workshop —— 119

Part III **Existence, Nonexistence, and Contradiction**

Friederike Moltmann
Existence Predicates —— 153

Kevin Mulligan
Modes of Being and the Mind —— 183

Takashi Yagisawa
Imagining Fictional Characters —— 203

Graham Priest
Objects That Are Not Objects —— 217

Part IV **Composition, Organisms, and Persons**

Peter Simons
The Concept of Organism and Degrees of Composition —— 233

William Jaworski
Peter van Inwagen and the Hylomorphic Renaissance —— 247

Alfredo Tomasetta
Remnant-Persons: A Commonsense Defence of Animalism —— 265

Part V **Abstract Beings, Nominalism, and Infinity**

Francesco Orilia
Van Inwagen's Approach to Relations and the Theory of O-Roles —— 279

Andrea C. Bottani
Properties, Nominalisms and Things That Can Be Said —— 297

Christian Kanzian
Paraphrase: A (More or Less) Van Inwagenian Way toward (Moderate) Nominalism —— 315

Joanna Odrowąż-Sypniewska
The Problem of the Many: Supervaluation, Rough Sets and Faultless Disagreement —— 329

Øystein Linnebo and Stewart Shapiro
Realizability as a Kind of Truth-Making —— 351

Part VI **God, Theodicy, and the Best World**

Nicholas Rescher
Optimalism in Explaining the Nature of Things —— 367

Chris Daly
Van Inwagen on Testimony and Contingency —— 399

Eleonore Stump
The Problem of Evil and Atonement —— 413

Dean Zimmerman
Resisting Rowe's No-Best-World Argument for Atheism —— 443

Mirosław Szatkowski
Deficiencies of Gödel's Ontological Proof —— 469

Authors of Contributed Papers —— 477

Person Index —— 483

Subject Index —— 489

Acknowledgements

My thanks are extended to many individuals and institutions, without whom I would not have been able to prepare this volume for publication. Above all, I would like to express my sincere gratitude to all authors who accepted the invitation to contribute to the volume and showed great understanding and patience during the publication process. Most authors participated in the - pre-publication of this book - international conference *Quo Vadis, Metaphysics*? *Dedicated to Peter van Inwagen to commemorate his* 75^{th} *birthday*, which took place in Warsaw on September 26 - 29, 2017.

I would like to thank the National Science Center for the co-financing support of both the conference and the publication of this book, within the research grant OPUS 4, DEC-2012/07/B/HS1/01978. I also would like to thank the Institute of Philosophy of the University of Warsaw for financial support of the conference.

Finally, special thanks are due to two doctoral students of the University of Warsaw, Ziemowit Gowin and Thomas Zyglewicz, for their great commitment to the preparation and implementation of the above-mentioned conference.

Mirosław Szatkowski

https://doi.org/10.1515/9783110664812-202

The Declaration of the Establishment of the *International Society for Formal Ontology*

On the occasion of the international conference "**Quo Vadis, Metaphysics? Dedicated to Peter van Inwagen to commemorate his 75th birthday**", which took place in Warsaw from 26 to 29 September 2017, 41 philosophers representing 17 countries have decided to establish the *International Society for Formal Ontology* (ISFO) based in Poland.

* * *

The date of the establishment of ISFO is September 27, 2017.

* * *

The aims of the ISFO are:
- integration of areas of research dealing with formal ontology,
- support of research in formal ontology,
- promotion of applications of formal ontology,
- connections of formal ontology with related fields.

* * *

Members of the founding committee of the ISFO will draft its statute and register the Society in accordance with the applicable Polish legislation.

* * *

Peter van Inwagen was invited by the founding committee to be the honorary president of the ISFO.

* * *

On behalf of the founding members, the declaration is signed by:

Gila Sher
(University of California, San Diego, USA;
Former Editor in Chief of the Journal *Synthese*, Netherlands;
Editor of *The Journal of Philosophy*)

https://doi.org/10.1515/9783110664812-203

Hannes Leitgeb
(Chair and Head of the *Munich Center for Mathematical Philosophy*),

Miroslaw Szatkowski
(Warsaw University of Technology, Poland)

Attachment: List of the founding members of the *International Society for Formal Ontology*.

Warsaw, September 27, 2017.

The Founding Members of the International Society for Formal Ontology

1. **Neil Barton** (University of Vienna, Austria),
2. **Christoph Benzmüller** (Free University of Berlin, Germany),
3. **Tomasz Bigaj** (University of Warsaw, Poland),
4. **Andrea C. Bottani** (University of Bergamo, Italy),
5. **Cezary Cieśliński** (University of Warsaw, Poland),
6. **Daniel Cohnitz** (University of Utrecht, Netherlands),
7. **Fabrice Correia** (University of Neuchatel, Switzerland),
8. **Mario De Caro** (Roma Tre University, Italy),
9. **Kit Fine** (New York University, USA),
10. **Peter Forrest** (University of New England, Australia),
11. **Paweł Garbacz** (The John Paul II Catholic University of Lublin, Poland),
12. **Michał Głowala** (University of Wrocław, Poland),
13. **Leon Horsten** (University of Bristol, UK),
14. **Guido Imaguire** (Federal University of Rio de Janeiro, Brazil),
15. **William Jaworski** (Fordham University, USA,)
16. **Janusz Kaczmarek** (University of Łódź, Poland,)
17. **Christian Kanzian** (University of Innsbruck, Austria),
18. **Srećko Kovač** (University of Zagreb, Croatia),
19. **Hannes Leitgeb** (University of Munich, Germany),
20. **Øystein Linnebo** (University of Oslo, Norway),
21. **Anna-Sofia Maurin** (University of Gothenburg, Sweden),
22. **Uwe Meixner** (University of Augsburg, Germany),
23. **Christopher Menzel** (Texas A&M University, USA),
24. **Friederike Moltmann** (French National Centre for Scientific Research, France; New York University, USA),
25. **Kevin Mulligan** (University of Geneva, Switzerland),
26. **Thomas Müller** (University of Konstanz, Germany),
27. **Peter Øhrstrom** (Aalborg University, Denmark),
28. **Francesco Orilia** (University of Macerata, Italy),
29. **Tomasz Placek** (Jagiellonian University, Poland),
30. **Carl Posy** (The Hebrew University of Jerusalem, Israel),
31. **Graham Priest** (University of Melbourne, Australia),
32. **Jonathan Schaffer** (Rutgers University, New Brunswick, USA),
33. **Benjamin Schnieder** (University of Hamburg, Germany),
34. **Gila Sher** (University of California, San Diego, USA),
35. **Ted Sider** (Rutgers University, New Brunswick, USA),

36. **Peter Simons** (University of Dublin, Ireland),
37. **Mirosław Szatkowski** (Warsaw University of Technology, Poland),
38. **Kazimierz Trzęsicki** (University of Białystok, Poland),
39. **Achille Varzi** (Columbia University, USA),
40. **Takashi Yagisawa** (California State University, Northridge, USA),
41. **Dean Zimmerman** (Rutgers University, New Brunswick, USA).

Mirosław Szatkowski

Introduction

The papers collected in this volume are in honor of the American analytic philosopher Peter van Inwagen. He received his doctorate from the University of Rochester in 1969, where his dissertation was supervised by Richard Taylor and Keith Lehrer. He taught for twenty-four years at Syracuse University, and has been the O'Hara Professor of Philosophy at the University of Notre Dame since 1995 (emeritus from 2019). For several years he has been (concurrently with his Notre Dame appointment) Research Professor of Philosophy at Duke University.

Peter van Inwagen is one of the most celebrated figures in philosophy of our time. His main research interests are in metaphysics, but he has also made significant contributions to ethics, philosophy of religion, logic, philosophy of mind, and philosophy of language.

He is the author of nine books: *An Essay on Free Will* (Oxford: Clarendon Press, 1983), *Material Beings* (Cornell University Press, 1990), *Metaphysics* (Boulder: Westview Press, 1993), *God, Knowledge and Mystery. Essays in Philosophical Theology* (Cornell University Press, 1995), *The Possibility of Resurrection and other Essays in Christian Apologetics* (Boulder: Westview Press, 1997), *Ontology, Identity, and Modality* (Cambridge: Cambridge University Press, 2001), *The Problem of Evil* (Oxford: Oxford University Press, 2006), *Existence* (Cambridge: Cambridge University Press, 2014) and *Thinking about Free Will* (Cambridge: Cambridge University Press, 2017); as well as the editor of five collections of original essays: *Time and Cause*: *Essays Presented to Richard Taylor* (Springer, 1980), *Alvin Plantinga* (with J. Tomberlin, D. Reidel Publishing Company, 1985), *Metaphysics*: *The Big Questions* (with Dean W. Zimmerman, Malden, Mass.: Blackwell Publishers, 1998), *Christian Faith and the Problm of Evil* (Grand Rapids MI: William B. Eerdmans Publishing Co., 2004), and *Persons*: *Human and Divine* (with Dean W. Zimmerman, Oxford: Oxford University Press, 2007). His publications include more than 200 articles in edited volumes and in professional journals, among them *The Journal of Philosophy*, *The Philosophical Review*, *Analysis*, *The Australasian Journal of Philosophy*, *Noûs*, *Philosophical Studies*, *The Monist*, *Philosophical Perspectives*, *The Philosophical Quarterly*, *American Philosophical Quarterly*, *Philosophical Topics*, *Ratio*, *Synthese*, and *Theoria*.

He has held many positions of distinction in professional societies, including member of the American Academy of Arts and Sciences since 2005, President of the Central Division of the American Philosophical Association in 2008/09, President of the Society of Christian Philosophers from 2010 to 2013, and Honorary President of the International Society for Formal Ontology since 2017. He is

https://doi.org/10.1515/9783110664812-001

also a present or past member of the editorial boards of many journals, among them *Faith and Philosophy*, *Noûs*, *Philosophical Perspectives*, *Philosophical Studies*, *Philosophy and Phenomenological Research* (to 2003), *Topics in Contemporary Philosophy*, and *Oxford Studies in Metaphysics*.

Professor Van Inwagen is an extremely sought after speaker by organizers of various conferences and congresses. He has given many conference and congress talks, and various lectures in the United States, Canada, Peru, the United Kingdom, Germany, Poland, and China. For example, some of his lectures include: the *F. D. Maurice Lectures* at the University of London, March 1999; the *Wilde Lectures on Natural Religion* at Oxford University in Trinity Term, 2000; the *Stewart Lectures* at Princeton University, October 2002; the *Gifford Lectures* at the University of St. Andrews, May 2003; the *Jellema Lectures* at Calvin College, March 2004; the *Stanisław Kamiński Memorial Lectures* at the John Paul II Catholic University of Lublin in Poland, May 2008; the *Münster Lectures in Philosophy* at the University of Münster, November 2015.

It is impossible to summarize in a short essay Peter van Inwagen's contribution to philosophy. So let's just mention a few, probably the more important and interesting, features of Peter van Inwagen's work:

1. A large part of Peter van Inwagen's works is devoted to the relation between determinism and free will. Van Inwagen claims that determinism is incompatible with free will and moral responsibility, and concludes that determinism should be rejected. He introduced to the philosophical literature the terms 'compatibilism' with respect to free will and determinism – the view that free will is compatible with determinism, and 'incompatibilism'– in contrast to the first term, and played an important role in rehabilitating libertarianism with respect to free will – the view that free will is real and that determinism is false. Van Inwagen, however, honestly says that the problem of free will is a great mystery that defies our best efforts.
2. The question of material composition – In what circumstances can simple objects be parts of whole objects? – occupies a prominent place in Inwagen's writings. Van Inwagen takes an intermediate position between two extreme positions: nihilism - the view that nothing is ever composed of anything else, and universalism - the view that any things compose another thing. According to him, some individuals – paradigmatically living things – are composed of physical materials with a specific organization or structure. One way to show that things are composed is: there is a y such that the xs composes y if the activity of the xs constitutes a life (or there is only one of the xs).

3. Peter van Inwagen is well known for his groundbreaking work on identity and personal identity – in particular, existence and existence over time – in particular, commonsense belief, modality, causality, the phenomenon of vagueness, and the relation between metaphysics and ordinary language.
4. Peter van Inwagen offered a new look on the problem of evil, namely, on the problem of answering arguments against the existence of God based on facts about evil. Van Inwagen rejects the thesis that there would be no evil in the world if there were a God. His answer is a quite complex defense of free will.
5. Finally, Van Inwagen's contributions should be mentioned on such topics as the problem of God's existence, the possibility of Resurrection and the Christian doctrine of the Trinity. Van Inwagen claims that each of the traditional arguments for the existence of God is – for different reasons – defective.[1] He – a philosopher who became a Christian at the age of forty – argues that God is capable of resurrecting the dead in some way or another. Regarding the Trinity, the doctrine as the proposition that there are three persons each of whom is God but just one being which is God, Van Inwagen's thesis is: "All the constituent propositions of the doctrine of the Trinity can be expressed in the language of relative identity, and they can be shown to be mutually consistent, given that the correct logic of identity is the logic of relative identity."

The present book includes papers from 17 speakers of the International Conference, dedicated to Peter van Inwagen to commemorate his 75^{th} birthday: *Quo Vadis, Metaphysics*?, Warsaw, September 26–29, 2017. This papers have been elaborated on the basis of delivered talks, taking into account the suggestions raised in the discussions. In addition, five texts – corresponding to the subject of the conference – were included, namely, by Professors Daly, Priest, Rescher, Stump, and Szatkowski. Both the title of the conference and this volume, on the one hand, with their dedications to Peter van Inwagen, on the other hand, were chosen very carefully and deliberately. The title *Quo Vadis, Metaphysics*? alludes to logical, epistemological and methodological considerations about the significance of metaphysics in all its richness, to reflections on its future perspectives, and – optimistically – to set the directions of its further development. One may wonder why the old philosophical discipline of metaphysics – after having been pronounced dead by many – has enjoyed a significant revival within the last thirty years. Without a doubt, Peter van Inwagen has significantly contributed to this revival. The attractiveness of the title of the conference and this volume, and the academic authority of Peter van Inwagen, together taken, produced a great deal of

1 Cf. "An Interview with Peter van Inwagen", this volume, pp. 7–28.

interest in the conference and the publication of the article in the post-conference book. Conference participants offered remarks such as the following: "This is an excellent and worthy initiative ... I would like to join in celebrating Peter's 75^{th} birthday", "It's a very nice idea to commemorate Peter van Inwagen's 75^{th} birthday with a conference on metaphysics in Warsaw", "I admire Van Inwagen's work and it will be an honor for me to participate in the conference", "Prof. van Inwagen is one of the few top metaphisicians in the world and it will be a great honor and a pleasure to discuss metaphysical issues with him and other distinguished philosophers in Warsaw", "It's great that you're doing a fest-conference for Peter van Inwagen".

The main content of the volume is divided into six parts: Part I. *Peter van Inwagen – A Profile*; Part II. *The Grounds and Ways of Metaphysics*; Part III. *Existence, Nonexistence, and Contradiction*; Part IV. *Composition, Organisms, and Persons*; Part V. *Abstract Beings, Nominalism, and Infinity*; and Part VI. *God, Theodicy, and the Best World*.

Part I includes Mirosław Szatkowski's "Interview with Peter van Inwagen", and Elisabeth M. Bolduc's short essay "Jonathan and Peter" on a friendship between the British philosopher Jonathan E. Lowe and Peter van Inwagen. Van Inwagen's answers to questions of the interview are an expression of his fascination with metaphysics and his longtime struggle with metaphysical questions. They sum up, in a very condensed way, his views on metaphysics, in particular – on its method, scope, specificity, future directions of research, but also his views on the value of traditional proofs for the existence of God. Van Inwagen very modestly speaks about his contribution to the building called 'metaphysics'.[2] According to Elisabeth M. Bolduc – Peter van Inwagen's wife – the love of metaphysics, the pursuit of truth, and naturally the sharing of common values – personal character traits were keystones of van Inwagen's friendship with Jonathan. Jonathan's untimely death in 2014 has given a different dimension to this friendship – the memory of Jonathan is for Peter like a hidden treasure, which he unearths in special hours.

Part II includes five essays. The first two of them, namely, "Where Are You Going, Metaphysics, and How Are You Getting There? – Grounding Theory as a Case Study" by Gila Sher, and "On the Relevance of Grounds" by Benjamin Schnieder, focuse upon grounding. Recent years have brought a surge in the metaphysical literature on grounding, generally speaking, on the relation that connects more and less fundamental entities. However, this term is vague, because there is no

2 We encourage the reader to read also Inwagen's answers to five questions posed in the book *Metaphysics*: *5 Questions* by Asbjorn Steglich-Petersen (ed.), Automatic Press (2010).

undisputable demarcation line between 'grounding' and other related notions, for example, such as 'causation', 'metaphysical causation', 'dependence', 'metaphysical dependence', 'explanation', 'metaphysical explanation', 'fundamentality' and 'metaphysical fundamentality'. In discussions about grounding two opposing tendencies are revealed – for some philosophers, 'grounding' is a primitive term, i.e., no helpful analysis of grounding in other terms can be given; for other philosophers, it is an analysable notion in other terms. According to many, grounding can and should be helpful in investigations of other interesting philosophical questions. Sher's paper proposes a new methodology, called *foundational holism*, in the approach to grounding. In turn, Schnieder's essay shows how the theory of grounding can be applicable to traditional philosophical debates about the Principle of Sufficient Reason, the question of why there is something rather than nothing, and the question of whether free will is compatible with determinism. Grounding sheds new light on these issues. The question of unity or multiplicity of metaphysics stands at the center of the third essay, "Metaphysical Differences" by Uwe Meixner. The pervasive and perennial conflict among metaphysicians seem to argue for the multiplicity of metaphysics. Meixner finds, however, arguments for the unity of metaphysics. The fourth essay, "A Van Inwagenian Defense of Constitutionalism" by Anna-Sofia Maurin, addresses two questions: (i) What is metaphysics?; and (ii) Do properties exist independently or depending on objects? As regards (i), Maurin agrees with Van Inwagen's view that there is no sense to say that metaphysics deals with explaining the nature and the existence of mind-independent reality. In reference to (ii), Van Inwagen argues that properties exist apart from the objects they characterize. Therefore, there exists no sense to say that properties 'make up' or constitute objects. Maurin, on the other hand, thinks: "it makes sense to say (although it may be false) that properties 'make up' objects, that properties exist 'in' the objects they constitute, and that properties are immanent universals or tropes." Finally, the fifth essay, "Inside the Metaphysical Workshop" by C. J. Posy, concerns the limits of applicability of formal semantics in metaphysics. The view that formal semantics – understood as the abstract account of truth (or, more generally, as the abstract theory of meaning) – has a central position in metaphysics, and that all metaphysical issues rest upon semantic presuppositions, is viewed as unacceptable.

Part III contains four essays. These essays can be organized around the following two questions: *What does it mean to ask if existence is a property of objects*?, and *Assuming that existence is a property of objects, are there objects that lack it*? The full answers to these questions require a broad and in-depth study, and certainly go beyond the scope of any four articles. Turning to specific essays, Friederike Moltmann in "Existence Predicates" discusses the wide rank of existence predicates functioning in natural language from the perspective of 'nat-

ural language ontology'. In his essay, "Modes of Being and the Mind", Kevin Mulligan discusses different modes of being (existence) admitted by Brentano and his heirs, and then he argues that different modes of being can be correlated with different types of mental acts. In turn, Takashi Yagisawa in "Imagining Fictional Characters" argues for the thesis that fictional objects are not abstract objects manufactured by people. His core thesis is that abstract objects, unlike fictional objects, are not subject to imagination. Finally, the essay "Objects That Are Not Objects", authored by Graham Priest, explores the issue *What is an object?*; otherwise, *What distinguishes objects from non-objects?* Unfortunately, there is no standard account how to classify entities as an 'object' or a 'non-object'. Philosophers have given many different answers to these questions that are extremely controversial and, consequently, they cause skepticism as to their usefulness. For a number of philosophers (for example, Wittgenstein, Heidegger, Frege) there are certain entities which appear to be both objects and not objects. Therefore, the question arises: Can something be both an object and not an object? Priest examines this question and proposes an answer.

Part IV contains three essays. Although there is no explicit structure to group these three essays, the first of them – "The Concept of Organism and Degrees of Composition" by Peter Simons – encompasses the issues addressed by the other two. Simons organizes his essay around the ontological problem of composition: *When and under what conditions do things combine so as to constitute or compose another thing?* Most generally, the question is: *When does composition occur?* There have been a number of responses to this question, for example, universalists say "always", while nihilists say "never". Simons analyzes these two extreme positions, but also the intermediate positions between them. But philosophers are not the only ones concerned with "composition" – scientists of various kinds (specifically, biologists, chemists and physicists) also study this phenomenon. According to Simons, comparing these two different approaches to composition, namely, metaphysical approach and scientific approach, can or should lead to exploring the mystery of "composition". In his essay, "Peter van Inwagen and the Hylomorphic Renaissance", William Jaworski explains how Van Inwagen's works on composition provide the basis for a hylomorphic theory that dovetails with current work in biology and its subdisciplines such as neuroscience. Finally, in "Remnant-Persons. A Commonsense Defence of Animalism", Alfredo Tomasetta seeks to explain the remnant-person problem put forward by Mark Johnston. This problem can be presented here briefly, in Tomasetta's words: "Your head detached from the rest of your body seems to be a person; now, if s/he is not you, then one can bring a person into being simply by removing tissue from something, and this is absurd. If, on the other hand, s/he is you, then animalism is false."

Part V includes five essays. In the first of this essays, "Van Inwagen's Approach to Relations and the Theory of O-Roles", Francesco Orilia analyses Van Inwagen's semantic problem, discussed in "Names for Relations", about names for non-symmetric relations and corresponding converses. This problem is articulated into two distinct ontological issues: (i) Whether there are converse relations, and (ii) Russell's problem of relational order – what accounts for the difference between two distinct facts or states of affairs such as Harry's loving Sally and Sally's loving Harry, which appear to involve the same relata and the same non-symmetric relation? By appealing to o-roles and by distinguishing sparse and abundant levels of properties and relations, Orilia tries to find solutions to these two problems. The next two essays - "Properties, Nominalisms and Things That Can Be Said" by Andrea C. Bottani, and "Paraphrase: A (More or Less) Van Inwagenian Way toward (Moderate) Nominalism" by Christian Kanzian – bring some new ideas to the debate between nominalism and Platonism. Bottani argues that the Van Inwagen's view that properties are things that can be said of things is far from being clear. And contrary to Van Inwagen, he gives such an interpretation of this view that does not entail Platonism about properties. Christian Kanzian endeavors to make sense of Van Inwagen's imperative: "We should wish not to be Platonists if it's rationally possible". In the fourth essay, "The Problem of the Many: Supervaluation, Rough Sets and Faultless Disagreement", Joanna Odrowąż-Sypniewska discusses three existing solutions to the problem of the many, namely, supervaluationism, fuzzy sets and Lewis supervaluation and almost-identity. In the last essay, "Realizability as a Kind of Truth-Making", Øystein Linnebo and Stewart Shapiro address the issue: How to make sense to universal generalizations in the context of potential infinity? In particular, can universal generalizations over all the natural numbers be true in the context in which only finitely many numbers can be generated? The authors use the intuitionistic notion of realizability to resolve this issue.

Finally, Part VI also includes five essays. The first essay, "Optimalism in Explaining the Nature of Things" by Nicholas Rescher, revolves around the Leibnizian question: Why is there anything at all? This question, according to Rescher, demands a collective explanation ("There is one single comprehensive explanation that accounts for all existents – the entire totality of them"), and not a distributive explanation ("There is some case-specific explanation to account for each and any individual existent"). Of particular importance here is what Leibniz says: "The reasons for the entire world [must] therefore lie in something extramundane, different from the chain of states or series of things whose aggregate constitutes the world. … So [to account for the world's being] there must exists something which is distinct from the plurality of beings, or from the world." Rescher adds that the ultimate explanation of of reality-as-a-whole can only be done in terms of value.

The crux of the reasoning required here lies in the *optimality principle*: "Given an exhaustive range of possible alternatives, it is the best of them that is actualized." This principle is self-explaining, that is, it does not require an external explanation in terms of something altogether different. Of course, there is a close relationship between the optimality principle and the question of determinism – the doctrine that the reality does not admit of any unrealized alternatives. The second essay, "Van Inwagen on Testimony and Contingency" by Chris Daly, assesses two of Van Inwagen's arguments concerning rational belief revision: the argument from expert testimony and the argument from contingency. The first argument concerns the topic of disagreement between equally informed and rational agents (epistemic peers). And the second argument concerns the contingency of our philosophical beliefs – we currently have some philosophical beliefs although we could have had others instead. In the third essay, "The Problem of Evil and Atonement", Eleonore Stump tries to build a theory for a unified approach to both the doctrine of the evil and the doctrine of atonement. She captures these doctrines in terms that are not primarily philosophical, but more theological. That means, she uses resources such as the Bible. Stump accepts the Christian teaching that the ultimate good for human beings is union with God, and that suffering has some role to play in promoting this ultimate good. In the fourth essay, "Resisting the No-Best-World Argument for Atheism", Dean Zimmerman critically evaluates the "no-best-world" argument: *If there is no best possible world for God to create, then God does not exist*. He rejects this argument in two ways: by denying that every creatable world could be improved upon, or by rejecting the normative principle that settling for a less-than-maximally-good world implies moral imperfection. The last essay, "Deficiencies of Gödel's Ontological Proof" by Mirosław Szatkowski, focuses on Godel's original ontological proof, who has received a lot of interest in the philosophical literature of the last 50 years. According to Szatkowski, Gödel's proof does not meet the basic deductive requirements, since: (1) Gödel's axioms do not result in a thesis for the existence of God; and (2) the thesis for the existence of God is inconsistent with Gödel's axioms.

Part I: **Peter van Inwagen – A Profile**

Mirosław Szatkowski and Peter van Inwagen

Interview with Peter van Inwagen

1

Mirosław Szatkowski:
Peter, you told me in one of your e-mails that metaphysics is not only the passion of your life, but something more – it is the love of your life.[1] How did you find this love? What were the beginnings of your fascination with metaphysics?

Peter van Inwagen:
When I was starting out in philosophy, when I was, so to speak, beginning to be a philosopher, I should have described my interests as centered not on "metaphysics" but on certain philosophical problems: the problem of free will and determinism, the problem of fictional existence, the nature of modality.[2] As time passed, however, I began to use the term 'metaphysics' to tie the members of this rather diverse set of problems together. (As I became interested in further problems – the nature of material objects and their relations to their parts, the problem of identity across time, the problem of nominalism and realism –, I continued to use the word 'metaphysics' as a general term to tie the problems I was interested in together. I do not think that I became interested in these further problems because someone had classified them as belonging to 'metaphysics'.) But why did I use *that* word? This is a hard question to answer because it is not at all clear what it

1 Peter van Inwagen added at a later date:

> I fear my wife and children will be distressed to discover that I have said that metaphysics is *the* love of my life, for they have been accustomed to regarding metaphysics (or accustomed to regarding philosophy – for they would not distinguish between metaphysics and philosophy) as no more than a serious rival for my affections. (If I am in my study, and my wife is asked where I am, she shrugs and replies, "Oh, he's with his mistress, Philosophia.")

2 My answer to this question and my answers to questions (2), (3), (5), and (5a) are "reprints" (with some adaptations and revisions – in most cases, minor ones) of my answers to some very similar questions in "Answers to Five Questions about Metaphysics," which was included in Asbjorn Steglich-Petersen (ed.) *Metaphysics*: *5 Questions* (Automatic Press/VIP, 2010). (See pp. 179–185.) My answers to questions (4), (6), (7), (8), and (9) appear here for the first time. (Perhaps in the case of questions (6) and (9), I should say 'responses' rather than 'answers'.)

https://doi.org/10.1515/9783110664812-002

means to classify a philosophical problem as metaphysical. I had long been aware that 'metaphysics' and 'metaphysical' were problematical terms, but I did not fully appreciate how problematical they were till a few years ago when I began to write the article "Metaphysics" for *The Stanford Encyclopedia of Philosophy*.

Even when I had not seriously thought about any other philosophical problem than the problem of free will and determinism, I described my interest in that problem as "metaphysical". (Or perhaps I said, "I'm interested in the *metaphysical* problem of free will and determinism" – implying that there was more than one philosophical problem that could be called 'the problem of free will and determinism' and that I was interested in the one that was metaphysical.) I said this because I believed that determinism – the thesis that only one future is consistent with the present state of things and the laws of nature (or the laws of physics) – was a metaphysical thesis and that any problem that essentially involved determinism was therefore a metaphysical problem.

But what did I mean by saying that determinism was a metaphysical thesis? That would be hard to say. I think it's clear what the, as one might say, phenomenology of my choosing that term was. Most other writers on the problem of free will and determinism did not think of determinism in the very abstract way that I did – or so at least it appeared to me. *They* were not thinking in terms of "the laws of nature" or "the laws of physics". *They* had not had scientific educations – not even the first few stages of a scientific education that I had had. *They* had never had to answer examination questions like, "An artillery piece is fired at an elevation of 37 degrees. The muzzle velocity of the shell is 2000 meters/second. What will the position and velocity of the shell be ten seconds later? (Neglect air resistance and the rotation of the earth.)." *I* could see that these examination questions had answers – as, of course, examination questions should. I could see that (neglecting air resistance and the rotation of the earth, to be sure), Newton's laws of motion and assumption that the acceleration due to gravity near the surface of the earth is a given that does not vary from case to case jointly implied that the elevation of a gun and the muzzle velocity of a shell fired from it were together sufficient to *determine* the position and velocity of the shell at any moment between the moment the gun was fired and the moment of impact.

Determinism, as I saw determinism, was a generalization of and abstraction from the fact that certain questions have answers – the questions about the evolution of physical systems that constitute such a high proportion of the exercises that one finds at the ends of the chapters in physics textbooks. (That is to say: the author of the text gives the student some numbers that describe the state of a system at one time and expects the student to produce some numbers that describe its state at some later time.) The generalization, however, and the abstraction are extreme, and their extremity takes one outside science. In making this generaliz-

ation one quantifies over laws of physics and the physical quantities that occur in them – over real laws of physics, God's-eye laws of physics, which may well be radically different from any of those principles that scientists and engineers of the present day use to grind out numbers that characterize the behavior of projectiles and planets and protons. And quantifying over real, God's-eye, laws of physics is not something that is done "within" the science of physics or within any other science. It was because my approach to the problem of free will and determinism had this sort of "feel" that I described it as 'metaphysical'. (As opposed to what? Well, as opposed to 'psychological', 'linguistic', 'commonsensical', 'ethical' – all words I used to describe the approaches to the problem of free will and determinism that I found in the work of various other writers.)

The preceding two paragraphs were an attempt to describe what was in my mind when I said that the determinism I was interested in was "metaphysical" determinism. (Other philosophers might use the word 'determinism' as a name for – say – the thesis that human action is determined to occur by the agent's desires and beliefs at the moment just prior to that action. That sort of thesis wasn't … well, metaphysical enough to engage *my* refined interest.) Perhaps this attempt was successful and perhaps not, but it was certainly not much help with the question, What did I *mean* by calling the kind of determinism I was interested in "metaphysical" determinism. After all, that question has an answer only insofar as I did mean something by 'metaphysical', and it's not now evident to me that there was anything much I meant by the word – or anything much beyond this: a philosophical thesis is metaphysical if (i) it can't be assigned with confidence to any other part of philosophy, and (ii) it involves a very high level of abstraction.

And what, if anything, do I mean by 'metaphysics' now? I have no interesting answer to this question. For an extended exploration of the question 'What does "metaphysics" mean?' (and for some difficulties I now see in an earlier attempt of mine to answer this question), see the article in *The Stanford Encyclopedia of Philosophy* that I mentioned above.

What keeps me interested in the questions I *call* metaphysical (beyond the interest each of them has for me *individually*, in and of *itself*: I just *am* interested in the problem of identity across time; I just *am* interested in the question whether there are abstract objects), is that the attempt to answer them seems in every case to involve a certain kind of thinking (there is a certain kind of thinking such that, in every case of a question I call metaphysical, when I attempt to answer that question I find myself engaging in that kind of thinking). It seems, moreover, that *only* the questions I call metaphysical call for that kind of thinking. I will attempt to describe the nature of this kind of thinking in my answer to question (5). Here I want to say something that is not about its nature but about what it is like to engage in it. I will do this by contrasting it with another kind of philosophical

thinking that I have some experience of. Most of my philosophical thinking that is not about metaphysics belongs to Christian apologetic. (Which does not of course imply that none of my apologetic thinking is metaphysical thinking – that would be false.) This thinking could be looked upon as being in the service of "applied philosophy". (When apologetic is done by a philosopher, it is generally fair to describe it as applied philosophy.) It is the kind of thinking one does when one is defending an ethical or political or aesthetic or religious position that one considers particularly important against some reasoned attack by an opponent of that position. A good example of the kind of thinking I have in mind can be found in my papers "Non Est Hick" and "Critical Studies of the New Testament and Users of the New Testament". If Christianity is not the illusion most philosophers suppose it to be, what I have done in these and other essays of the same type may well be – depending on how good it is and whom it has reached – more important, perhaps vastly more important, than my work in metaphysics. But it is clear to me from my own experience of engaging in the kind of thinking that goes into these essays that that thinking does not engage the full resources of my mind. And that is not what I would say of the kind of thinking on display, for good or ill, in *Material Beings* or the essays collected in *Ontology, Identity, and Modality* and *Existence: Essays in Ontology*. Only when I am thinking about matters like "the special composition question" or Lewis's modal ontology or Putnam's criticisms of Quine's ontological method do I feel that my mind is fully awake. (I do not identify myself with my mind; I am not saying that I am fully awake only when I am engaged in metaphysical thinking. One in fact doesn't want one's mind to be fully awake any very high proportion of the time – if for no other reason, because when one's mind is fully awake, one's capacities for interacting with other human beings in all sorts of important ways will be asleep. If the Good Samaritan's mind had been fully awake when he was on the road from Jerusalem to Jericho, he would have been too wrapped up in his own thoughts even to have noticed the man who had fallen among thieves.) And this sort of thinking is addictive. I hope that when I am no longer able to do it, I shall be aware of this fact and able gracefully to stop trying to it. Till then, however, I have no choice but to continue indulging my addiction.

Having re-read what I have just written, it occurs to me that it may well be that I call a question metaphysical just in the case that my attempt to answer it involves the kind of thinking I have been trying to describe.

2

Mirosław Szatkowski:
You are the author of 6 books and over 200 articles. How would you characterize your work? Which problems have especially absorbed you? What answers have you given. What is your contribution to the edifice of metaphysics?

Peter van Inwagen:
I think I did as much as anyone to undermine the view that was the consensus on the problem of free will in the middle sixties when I began graduate studies in philosophy. This view was that the problem of free will was a solved problem. And the solution was 'compatibilism': the thesis that free will and determinism are compatible (because 'X was able to do otherwise' means something conditional, something along the general lines of, 'X would have done otherwise if X had chosen to do otherwise').

I think also that I left the problem of free will and determinism *clearer*, more precisely stated, than I found it. (It saddens me that those now working on the problem of free will and determinism are, or a significant proportion of them are, engaged in simply throwing all that hard-won clarity away. If one examines a really clear piece of writing on the problem of free will and determinism – for example, David Lewis's great essay, "Are We Free to Break the Laws?" – and the kind of thing that makes up no small part of what is written about free will and determinism today, the contrast is astonishing.)

I attach some importance to my defense of an "abstractionist" modal ontology – and particularly to my reply to David Lewis's charge that anyone who claims so much as to *understand* the language in which abstractionists frame their modal ontology is in effect claiming to possess magical powers of understanding.

I think that I did as much as anyone to create "the problem of material constitution". And I was certainly the philosopher who brought the "Special Composition Question" to the attention of the philosophers who were working on material constitution (despite the fact that I was not the first philosopher to formulate the question).

I think I have had some important things to say certain problems about the identity of things and persons across time. I think that some of the things I have said about the concept of a temporal part and about the psychological-continuity theory of personal identity are worth paying attention to.

I believe I am responsible for metaphysicians' having come to think in terms of a distinction between 'ontology' and 'meta-ontology' – ontology being the discip-

line that asks the question 'What is there?' and meta-ontology being the discipline that asks the question, "What are we asking when we ask 'What is there?'?"

3

Mirosław Szatkowski:
Which metaphysical issues remain open? What areas of research would you recommend to future generations?

Peter van Inwagen:
At the turn of the millennium, I should have recommended that metaphysicians pay serious attention to the field I dubbed 'meta-ontology'. Happily, no such recommendation is now necessary. I hope that the current lively debates about meta-ontology (such as those on display in the 2009 collection *Metametaphysics*: *New Essays on the Foundations of Ontology*) will continue and deepen.

I hope that in the coming decades, metaphysicians will devote considerably more time than they so far have to the topic of the relative merits of constituent and relational ontologies.

Constituent ontologies are ontologies that affirm the existence of attributes (properties, qualities, characteristics, features) and which, moreover, treat these objects as being in some sense "constituents" of the substances (individuals, particulars) that have them (exemplify them, instantiate them, exhibit them). The theory that individuals are "bundles" of qualities is a paradigmatic example of a constituent ontology – for if *x* is a bundle of *y*s, those *y*s must in some sense be constituents of *x*.

Relational ontologies are ontologies that affirm the existence of attributes but which treat the "having" relation as in no way like the whole-to-part relation – as not even remotely analogous to that relation or to any mereological relation. According to the advocates of relational ontology, the binary relation "having" that Mars and a socialist banner bear to the quality redness is as abstract, as bloodless, as purely external, as the variably polyadic relation "are numbered by" that the moons of Mars and the epics of Homer bear to the number two. It is an axiom of relational ontology that the only "constituents" of any substance (individual, particular) are its parts, its parts in the strict and mereological sense; and, further, that any proper parts a substance (individual, particular) has are "smaller" members of the same ontological category: smaller substances (smaller individuals, smaller particulars).

Many metaphysicians have endorsed and have worked within a constituent ontology. Many metaphysicians have endorsed and have worked within a relational ontology. But an examination of the relative merits of constituent ontologies (on the one hand) and relational ontologies (on the other) is a neglected and important topic.

I also hope that some metaphysicians will turn their attention to the question of the implications a relational ontology for the philosophy of mind. My own investigation of this question can be found in, "A Materialist Ontology of the Human Person" in the collection *Persons*: *Human and Divine* (Oxford: 2007) that Dean Zimmerman and I edited, and in "Causation and the Mental", in *Existence*: *Essays in Ontology*.

4

Mirosław Szatkowski:
Don't you think that metaphysics is a 'machine' producing questions to which there are no answers whatsoever, or to which there is a multitude of competing answers, none of which can be selected as unambiguously correct?

Peter van Inwagen:
I don't know what to do with the "machine" metaphor. But I will say this. Some of the interrogative sentences put forward by metaphysicians as supposed vehicles of metaphysical questions are meaningless – and are thus not vehicles of questions at all – and some are not. Metaphysical questions, of course, have answers – for they are *questions*, and, as Wittgenstein has said, "The *riddle* does not exist. If a question can be put at all, it can be answered."[3] (I interpret 'it can be answered' as 'it has an answer', and not as 'it is possible for someone to answer it'.) But I doubt whether our human minds – in our fallen condition, in this present life – are so constituted as to enable us to answer very many of them (if any). (Compare this remark to what I say about philosophical arguments in my answer to Question 8.) Of course, if one metaphysician answers the question "Are there universals?" Yes and another answers it No, one of them must be right. (Assume for the sake of the example that the sentence 'There are universals' is meaningful and, moreover,

3 *Das R ä t s e l gibt es nicht. Wenn sich eine Frage überhaupt stellen läßt, so kann sie auch beantwortet werden.* – *Tractatus*, 6.51

has a precise enough meaning for it to express one determinate proposition.) But, I contend, the answer given by whichever of them is right will almost certainly lack *warrant* (in Plantinga's sense).

I must add this to what I have said: the same is true of questions posed in moral and political philosophy, in epistemology, in the philosophy of mathematics, and in all other parts or areas of philosophy. Metaphysics is no worse off in this respect than any of those other parts of philosophy. And that includes metaphilosophy and therefore includes the critiques of metaphysics offered by philosophers like Kant (well, in his case, his critique of *transcendent* metaphysics) and Carnap and van Fraassen. A metametaphysical proposition like 'Metaphysical statements are meaningless' is in the same unsatisfactory epistemological position as a metaphysical proposition like 'Qualities inhere in partarticulars' or 'There is more than one mode of being'. And a metametametaphysical proposition like the one expressed by the previous sentence is in that position, too – and so *ad infinitum*.

5

Mirosław Szatkowski:
Which research method in metaphysics would you favour? Does such a thing as specific metaphysical cognition exist?[4]

Peter van Inwagen:
William James has said, "Metaphysics means only an unusually obstinate attempt to think clearly and consistently." While this will hardly do as a *definition* of metaphysics, it is not a bad statement of the only *method* we metaphysicians have. A fuller attempt to answer this question can only take the form of a series of footnotes to this statement – can only be an attempt at a statement of what a metaphysician's obstinate attempt to think clearly and consistently should involve.

Bas van Fraassen, an avowed enemy of metaphysics, seems to believe that the method of metaphysics (insofar as a pseudo-discipline can have a method) is that of "inference to the best explanation". As scientists are said by some to

4 The original form of question (5) was "Which research method in metaphysics would you favour? Does such a thing as specific metaphysical cognition exist? In what way can it be valuable to other research disciplines?" I have taken it upon myself to "split" this question into (5) – as it appears in the text – and (5a), which is "In what way can metaphysics be valuable to other research disciplines?"

survey a set of empirical data and then try to come up with a theory that is the best explanation of those data, metaphysicians, van Fraassen maintains, (think they) proceed by surveying some set of data (I will not attempt to say what these data might be) and then attempting to construct theories that explain them. These metaphysicians (so they suppose) then proceed to compare the theories they have constructed to explain one of these sets of data with an eye to discovering which one best explains them. (What the standards of comparison are, I will not attempt to say.) And it may be that van Fraassen is right to say that this is what some metaphysicians (think they) are up to – and right in his unflattering comparison of the fruits of their labors with those of the labors of physicists and geologists and microbiologists. Van Fraassen errs, however, in supposing that this "method" (I agree entirely with his low opinion of its fruits) is essential to metaphysics, and I am doubtful whether it is very commonly employed by philosophers who call themselves metaphysicians. Like many people who offer unflattering diagnoses of the ills that afflict some field of human endeavor, van Fraassen has fallen in love with his diagnosis and applies it indiscriminately and uncritically. "You're one of the people he's applied it to, right?" Very perceptive, Reader. But if I use my own work as an example, at least I'm in a position to have an informed opinion concerning the method of the person I'm using as an example. Van Fraassen has written,

> When interpreting scientific theories, we see might careful attention to the empirical aspect, and the relation of the empirically superfluous parameters introduced to the observable phenomenon. That is why the Cartesian theory of vortices should receive considerably more respect – I'll say the same about Bohm's particles – than, e.g., Peter van Inwagen or David Lewis's mereological atoms. Mere observance of correct logical form does not make a theory genuinely valuable: in Tom Stoppard's phrase, it can be coherent nonsense. ("Replies to Discussion on *The Empirical Stance*", *Philosophical Studies*, Vol. 121/2 (2004), pp. 171–192. The quoted passage is on p. 181.)

If I understand what van Fraassen is saying, he thinks that the "mereological atoms" that occur in a certain metaphysical theory of mine – the theory presented in *Material Beings* – are "there" for some metaphysical reason: that they are a "metaphysical posit", that I have *postulated* them because, in my view, postulating them aids in explaining some set of data I have set out to explain. In fact, however, the mereological atoms are there because, rightly or wrongly (wrongly, Ladyman, *et al*. would say), I thought that the physicists said that matter had an atomic structure. Feynman has said:

> If in some cataclysm all scientific knowledge were to be destroyed and only one sentence passed on to the next generation of creatures, what statement would contain the most information in the fewest words? I believe it is the atomic hypothesis (or atomic fact, or

> whatever you wish to call it) that all things are made of atoms – little particles that move around in perpetual motion, attracting each other when they are a little distance apart, but repelling upon being squeezed into one another. In that one sentence, you will see there is an enormous amount of information about the world, if just a little imagination and thinking are applied. – Richard Feynman, Robert B. Leighton, and Matthew Sands, *The Feynman Lectures on Physics*, 3 Vols. (Reading, Mass.: Addison-Wesley, 1963–65), Vol. I, p. 2.

Feynman, of course, is talking about atoms in the modern, chemical sense. In that sense, "atoms" are not what van Fraassen calls *mereological* atoms – but Feynman would certainly not have objected to the statement that, just as "all things" (all things that are present to the senses or that can be seen through an optical microscope) are "made of (chemical) atoms," so chemical atoms are made of electrons and protons and neutrons (and perhaps photons), and protons and neutrons are made of quarks (and perhaps gluons). And there are good empirical reasons to suppose that electrons and quarks (and photons and gluons) are not "made of" anything (or, if you like, that they are not represented by the "standard theory" of elementary particles as made of anything): that they are (represented as) mereological atoms. And there is good reason to think that future physical theories, successors to the standard theory, if they do not postulate electrons and so on, will postulate partless things (little vibrating "loops of string," perhaps – but little loops of string that neither have proper parts nor are made of a stuff called string.) It is as certain as anything in this area can be that no physics descended from present-day physics is going to represent the physical world as consisting of continuous, homeomerous Aristotelian matter or as consisting of "gunk". Physics is (pretty clearly) always going to be "atomistic" in *some* not entirely empty sense. Physics is always going to have to find some sense for statements like, "The matter – the *stuff* – that was in this test tube after the reaction is the same matter that was in it before the reaction – albeit in a different form." And this sense, when spelled out, is (pretty clearly) always going to involve phrases of the form 'same Xs' where 'Xs' represents a plural count-noun. So what I am I *supposed* to do when I'm constructing a metaphysical theory about the identities of physical objects across time – a theory that involves the notion of "same matter"? Adopt an *Aristotelian* understanding of "same matter"? No, I simply borrowed the current scientific account of "same matter" (and perhaps registered my conviction that any future scientific account of "same matter" will be like the present-day account in being – in a very broad sense – atomistic). In sum, the mereological atoms are present in my metaphysical theory simply because I believe what the physicists tell me about matter – or at any rate, I believe what I *believe* they've told me. Even if I've misinterpreted them, even if my understanding of them is as feeble as Ladyman *et al*. think it is, my mereological atoms are not present in my

metaphysical theory for a metaphysical reason. Van Fraassen thinks that they are only because he has brought to his reading of *Material Beings* a theory about what metaphysicians think they are doing – a theory that tells him that that's what I'm doing.

Whether or not this is fair to van Fraassen, I do, as I have said, agree with his contention that trying to construct theories that explain some set of data is not going to yield any metaphysical conclusions of any interest. But then what method or methods *should* metaphysicians employ? I would not presume to dictate to other metaphysicians how they ought to proceed – or not beyond urging them to make an unusually obstinate effort to think clearly and consistently. But I'll say a few things about what I try to do when I'm doing (what I call) metaphysics.

First, in metaphysics (and I would say, in all parts of "core" philosophy: metaphysics, epistemology, the philosophy of language, the philosophy of mind, and philosophical logic) all words and phrases should be used in their ordinary senses or else explicitly defined. (Physics textbooks can provide some very instructive examples of good, precise definitions of technical terms. I expand on this point in my answer to question (5a).) Definitions should satisfy the following formal requirement. They should be in "Chisholm style": the *definiendum* should be a sentence – normally an open sentence or a sentence schema – and the *definiens* a sentence containing the same free variables or schematic letters. In metaphysics, all terms of art should be connected to ordinary language by a chain of Chisholm-style definitions. (What I mean is that such a chain of definitions should be possible in principle, implicit in one's text and easily extracted from the text. One may certainly introduce one's terms of art more informally if one is confident that the reader will be able to see how to construct the chain of definitions. There's no call for unnecessary formality. But in borderline cases it's always better to err on the side of pedantry – for recall Russell's definition of a pedant: 'A man who cares whether what he says is true'.) Similarly, one's arguments should be formally valid – though not necessarily presented in a form that is explicitly so. To say this is not to imply that there are proofs in philosophy as there are proofs in mathematics. It is simply to recommend a trick that will ensure that one is at least *aware* of all one's premises.

While we are on formal matters, I insist that in core philosophy one be scrupulous about use and mention. Every metaphysician must understand "Quine Corners" or "quasi-quotation marks" and use them when they are appropriate. (In my experience, about eighty per cent of the philosophers who use Quine Corners use them impressionistically, without actually understanding how they work.)

Following these simple rules will enable the philosopher at least to produce what van Fraassen has called coherent nonsense. In my view, it's much better to write coherent nonsense than to write incoherent nonsense. The reason is simple:

if nonsense is logically coherent, it's much easier to see that it's nonsense and to see why it's nonsense than it is if the nonsense is logically incoherent. For example, if a philosopher's sentence contains a gross use-mention confusion, a reader of the text in which it occurs may suspect that there is some meaningful thesis that the author was *trying* to express – and may find, after re-writing or attempting to re-write the sentence without the use-mention confusion, that there was really no idea there at all. If the author had taken the trouble to write coherent nonsense, the reader would have been spared that task.

But these matters – important though they are – are of merely formal significance. What can I say that is more substantive? I would say that my own method in metaphysics (insofar as I have one) is this:

> One should consider those theses that one *brings* to philosophy – theses that (so one supposes) practically everyone, oneself included, accepts, or theses (so one supposes) that have been endorsed by disciplines other than philosophy and in which one reposes a high degree of confidence (economic history, it may be, or microbiology or algebraic topology). One should try to discover what the metaphysical implications of those theses are. If, for example, one wants to know whether there are universals, what one should *not* do is this: collect a set of data ("This thing here is red and that other thing over there is also red") and attempt to discover whether those data a best explained by a "theory" that "posits" universals; what one *should* do is to ask whether the theses that one brings to philosophy *logically imply* the existence of universals (one will, of course, have provided a careful definition of 'universal').

Note that this "method" (better: this piece of methodological advice) has implications for the epistemology of metaphysics. It implies the epistemological problems or questions that confront metaphysicians – those of them who employ this method – fall into two groups: questions that are raised by the things they believed before they came to metaphysics, and questions that are raised by their beliefs concerning the logical implications of those things. (For example: How can one determine whether the existence of the real numbers is a logical implication of the statement that there are bodies whose behavior is governed by the law of universal gravitation?) The questions in the first group are profoundly difficult, but they are not questions that confront metaphysicians because they are metaphysicians: they confront metaphysicians only because, outside or prior to philosophy, they believe what most people believe. (Obviously, therefore, the metaphysician who employs this method will be, in Strawson's words, a "descriptive" rather than a "revisionary" metaphysician.) The questions in the second group are no doubt difficult – some of them are difficult –, but there does not seem to be any good reason to regard them as intractable.

It is important to realize that I have *not* recommended the following method: Treat the theses we accept before we come to metaphysics as data that it is the

business of metaphysics to explain; construct metaphysical theories that explain those data; compare these theories and find the one among them that best explains those data. (The so-called Quine-Putnam Indispensability Argument is an example of this method at work.) No, I'm recommending only that metaphysicians try to discover the metaphysical implications of – the metaphysical theses that are *logical* implications of – the things they believe on non-metaphysical (and, more generally, non-philosophical) grounds.

There is another method, or another methodological idea, that has, I believe, profoundly influenced my own work. But I find this "idea" very difficult to formulate verbally. My best attempt is along these lines:

> Let your investigations be centered on general theses, not particular examples. If an otherwise attractive general thesis seems to have counterexamples, try to explain them away. If it is in conflict with particular things we are inclined to say, try to explain the fact that we are inclined to say these things away. Look at the particular theses about things in the light of the general theses you find attractive.

This methodological idea played a central role in the development of the theory I presented in *Material Beings*. In that case, it took something like this form:

> Do not begin your investigation of the metaphysics of material objects by asking, e.g., whether there are tables of chairs. Begin by considering possible alternative answers to the Special Composition Question. If the best answer seems to be one that implies that there are no tables or chairs, try to explain the fact that "We all think there are tables and chairs" away. Ask yourself whether there really *is* such a fact as this.

But *Material Beings* is a special and very difficult case. (Many philosophers believe the book to be an essay in revisionary metaphysics. And many who are not guilty of that misreading would be hard-pressed to find a way to regard it as an example of "trying to discover the metaphysical implications of things we all believe". I do so regard the book, but I cannot defend this view here.) Instead I will give a relatively simple example of the method I am recommending, an example drawn from philosophical logic rather than metaphysics. (It can be more briefly stated and raises fewer side issues than any example I can think of from metaphysics.)

The sentence-schema '$p \rightarrow (\neg p \rightarrow q)$' is a theorem of standard sentential logic. Many philosophers say that this fact implies that '$\rightarrow$' does not represent the 'if-then' of "ordinary" English conditionals ("'is'-'is'" conditionals, as opposed to "'were'/'did'-'would-be'" conditionals). If it did, they contend (the example, of course, is my own), the sentence

> If New York is not in the United States, New York is in California

would be true – which obviously it isn't. And how do they know that it isn't true? Well, they ask themselves whether it's true, and they discover within themselves a conviction that it isn't. In my view, the view embodied in the methodological principle I'm recommending, this isn't what they should be asking. They should, rather, be asking themselves what general logical principles they think govern 'if-then' (and 'or' [inclusive] and 'it is not the case that' and the other little English words and phrases that in some sense correspond to the connectives of formal sentential logic). I would ask these philosophers to consider the following two "ordinary language" logical principles:

> Addition: *p hence*, *p* or *q*
> Disjunction-Conditional: *p* or *q hence*, if it is not the case that *p*, then *q*

(or, in a rather more long-winded form: it is either the case that *p* or the case that *q*; *hence*, if it is not the case that *p*, it is the case that *q*).

And I would ask them to consider the following logical deduction:

1.	New York is in the United States	Premise
2.	Either New York is in the United States or New York is in California	(1), Addition
3.	If New York is not in the United States, then New York is in California	(2), Disjunction-Conditional

No one, I suppose, would be inclined to dispute the sole premise of this argument. And, therefore, if both Addition and Disjunction-Conditional are valid, our little deduction amounts to a *proof* of the conditional 'If New York is not in the United States, New York is in California'. One must, therefore, reject one of the two theses:

> Addition and Disjunction-Conditional are both valid,
> 'If New York is not in the United States, New York is in California' is not true.

The right choice seems to me to be evident: We must accept the first thesis and therefore must reject the second. We may say of the statements 'Addition is valid' and 'Disjunction-Conditional is valid' what Gödel said of the axioms of set theory: they *force themselves upon the mind as true*. (Who would reject the reasoning contained in this passage:

> There are two blackboards, *A* and *B*. Each has a single sentence written on it. The sentence written on *A* is true. Therefore, at least one of the two blackboards has a true sentence written on it?

Or in this:

> There are two blackboards, *A* and *B*. Each has a single sentence written on it. At least one of the two has a true sentence written on it. Therefore, if the sentence written on *A* is not true, the sentence written on *B* is true?[5])

As to the second thesis – well, who *cares* what truth-value that bizarre sentence has? – isn't its truth-value a *paradigm case* of a philosophical "don't care"?

As I recommend considering questions like, "What is the truth-value of the conditional 'If New York is not in the United States, New York is in California'?" only in the light provided by the consideration of much more general logical questions, so I recommend considering questions like, "Are there are tables and chairs?" only in the light provided by the consideration of much more general ontological questions.

5a

Mirosław Szatkowski:
In what way can metaphysics be valuable to other research disciplines?

Peter van Inwagen:
I think that philosophy in general, and metaphysics in particular, have very little to offer the natural sciences. (Philosophy and metaphysics are none the worse for that – just as sociology is none the worse for having nothing to offer to astrophysics.) In making this statement, I mean the phrase 'the natural sciences' to be understood in its strictest sense – I mean 'the natural sciences' to refer to the kind of research that leads to publications in journals of molecular biology or paleontology or condensed-matter physics. It is, however, a commonplace that not all scientists are content to communicate information about their work only in the pages of such journals – only to their peers, only to specialists in their own and closely related disciplines. According to Bouwsma, Wittgenstein once said (in conversation), "This is the age of popular science, and so cannot be the age of philosophy." I think that this characteristically gnomic statement means something like this: This is an age in which popular science plays a role in the general

5 Imagine someone who judges that that this reasoning is valid – and who repents that judgment when it is revealed that the sentence written on *A* is 'New York is in the United States'.

intellectual life of our species that had been played in an earlier age by philosophy (and in a still earlier age by theology). If this is true – and I think it is –, its truth is at least partly explained by two facts: that in the present age, scientists can expect that large numbers of people will listen to what they say on any subject they care to talk about, and that much of what appears under the rubric 'popular science' is, to all intents and purposes, philosophy. And this philosophy, the philosophy that infuses many works of popular science, is, I make bold to say, *radically amateur* philosophy, the philosophy of writers who do not know that there is such a thing as philosophy. (These writers no doubt know that there is something *called* 'philosophy' but they are unaware that this thing has any bearing on what they are trying to say – or perhaps a few of them do know that they are doing this thing called 'philosophy' but assume that, being scientists, they will automatically and without any resources beyond the furnishings of their own minds, be able to do it better than its official practitioners.) I have never seen any philosophical work by scientists (Galileo is the sole exception I am willing to allow) that is of much philosophical interest. And this judgment certainly applies to the attempts of scientists to discuss metaphysics. The attempts of scientists to address large questions outside their own disciplines (but informed by their knowledge of their disciplines) in work addressed to the general public would certainly be much better for some knowledge of what philosophers have had to say about those and related questions.

If metaphysics has nothing to offer the sciences, the sciences – the fruits of the real work of scientists and not their amateur attempts at philosophy – have a great deal to offer metaphysics. Many scientific discoveries are not only relevant to metaphysics but of inestimable metaphysical importance (one might cite the discovery by cosmology that the physical universe had a beginning in time, or the discovery by high-energy physics that material things are ultimately composed of things that are not themselves composed of smaller things – and yes, I know what McCall, Ladyman, and Ross have to say about that thesis). Nevertheless, the exploitation of this important resource for metaphysics (and more generally for philosophy) has been entirely the work of scientifically literate philosophers.

I would also note that, quite apart from the *discoveries* of physics (and the other sciences), many metaphysicians could learn a great deal by carefully studying the way in which the writers of physics textbooks introduce such concepts as "displacement," "velocity," "acceleration," "mass," "force," "momentum," "energy," "work," "power," and "heat."

I do wish that my colleagues in literature and the social sciences would stop trying to do metaphysics (well, it's generally *anti*-metaphysics that they're trying to do). The scientists, philosophical amateurs though they may be, at least have at their disposal a fund of propositions that can serve as premises in metaphys-

ical arguments. The *littérateurs* and the social scientists, however, have no such fund on which to draw. (I do think that literature and the social sciences are relevant to philosophy – and indeed of great philosophical importance. But their relevance is to ethics and political philosophy and, of course, to aesthetics, and not to metaphysics or anti-metaphysics.)

It has long been a complaint of mine that the philosophy of mind suffers from the failure of philosophers of mind to pay sufficient attention to the metaphysical issues their statements involve them in. When I try to read through – as an interested outsider – the course of various debates in the philosophy of mind, I often find them difficult to follow. (That's the polite way of putting my point. The less polite is: I constantly find myself saying, "What does that even *mean*?") In a typical work in the philosophy of mind, concepts – and more often than not, they're metaphysical concepts – are pulled out of the air with no attempt to provide them with any definition or analysis. I will provide two examples of what I'm talking about.

Philosophers of mind like to talk about 'states' – mental states, physical states, what-have-you states. And when you ask a philosopher of mind what a 'state' is, the reply is generally either a blank stare or something along the lines of, "Well, *you* know – *states*. Please, none of your metaphysician's ontological quibbling. We philosophers of mind know what we mean when we talk about mental states and physical states, and if you don't, that's *your* problem." I insist on ontological quibbling, however. I insist on asking whether a state is an *attribute* (or a *property*, *quality*, *characteristic*, or *feature*). These are abstract objects, things that exist in all possible worlds and which are without causal powers. And the answer to this question I insist on asking (when any answer is given) is usually something like, "Well of course that's not what states are. A person's mental states exist only when he or she is *in* them, and they're constantly causing and being caused by other states." And then I have to ask, "But what is there for a state to be but a property? Aunt Milly's mental and physical states aren't *substances*, are they? – that is, things that belong to the same ontological category as Aunt Milly herself?" It is rare for the conversation to get as far as this, but if it does, I'm told (I paraphrase), "Well, they're neither substances nor attributes, they're *states*. Don't expect the things we talk about in the philosophy of mind to fit into the neat *a priori* categories you metaphysicians dream up." And my rejoinder is, "I don't see any reason to believe that there are any things with the combination of properties you assign to 'states.' It looks to me as if the very idea of a thing that has those properties makes no sense. All the stuff you say about or in terms of 'states' looks to me as if it's *not even wrong*." (A closely related point: don't get me started on the radical ontological – and even logical – confusions that infect what philosophers of mind say when they start talking about "qualia.")

My second example is the psychological continuity theory of personal identity. But I have had a great deal to say about this subject already. I refer the interested reader to my essay, "Materialism and the Psychological Continuity Account of Personal Identity" (*Philosophical Perspectives*, Vol. 11: *Mind, Causation, and World* (1997), pp. 305–319).

6

Mirosław Szatkowski:
Let me quote one passage:

> Peter van Inwagen is one of the most eminent and influential Christian philosophers in the world today. His work in metaphysics and philosophy of religion is cited and taught very widely and he is a frequently invited speaker at conferences in many countries. He writes with clarity and power, and his views have helped to shape the thinking of a whole generation of philosophers. In all this work, his Christian convictions shine through. In an age when many are happy to dissemble either their adherence to Christianity or some part of their Christian convictions, Van Inwagen has been exemplary in his courage and his witness to his faith. He is widely honored for his intellectual gifts, but they are certainly matched by his integrity and his fidelity to the faith. Without doubt, he is an exemplar of what a Christian philosopher should be.

Peter van Inwagen:
I blush – and pray to be delivered from the spiritual danger in which reading such statements places a Christian philosopher or theologian. ("What doth it profit thee to enter into deep discussion concerning the Holy Trinity if thou lackest humility and art thus displeasing to the Trinity?" – Thomas à Kempis.)

7

Mirosław Szatkowski:
In your works you have also addressed the issues of the existence of God and His attributes. Do theoretical considerations influence your actions in life?

Peter van Inwagen:
Perhaps theoretical considerations in moral philosophy and revealed theology have influenced my "actions in life". I do not believe that theoretical considerations in metaphysics have – or at any rate, only in the sense that theoretical considerations in metaphysics may have helped me to see the flaws in certain widespread arguments for the falsity of various theological propositions that are essential to the Christian faith.

8

Mirosław Szatkowski:
How do you assess the value of proofs for the existence of God, especially the ontological ones or those by Thomas Aquinas? Why do you consider yourself an analytical philosopher instead of a Thomist? How do you assess Thomism – its advantages and disadvantages – from this external perspective?

Peter van Inwagen:
I will say first that I think that no philosophical argument whose conclusion is a positive, substantive proposition – and the proposition that God exists is a positive, substantive proposition if any proposition is – need convince those who perfectly understand it that its conclusion is true. I do believe that some versions of the cosmological argument are as good as any philosophical argument I've ever seen – and I would make almost as strong a statement about some versions of the design argument. But I am convinced that someone of high intelligence and unexceptionable intellectual honesty could understand these arguments perfectly and nevertheless reject their conclusions – and not thereby convict himself or herself of irrationality.

I would say, moreover, that none of the traditional "arguments for the existence of God" – valid or invalid, sound or unsound – really is an argument for the existence of God. Perhaps I can make the reason why I say this clear by considering the conclusion of the version of the cosmological argument that I like best; its conclusion is that there is at least one metaphysically necessary being, and that all metaphysically contingent beings depend for their existence on one or more of these necessary beings. And an atheist could accept that conclusion. (Not happily, perhaps, but certainly consistently.) Even the conclusions of the various versions of the ontological argument cannot really be read as affirmations of the existence of God. There exists *aliquid quo nihil maius cogitari possit*; well and good. But does

God exist? There exists *ens summe perfectum*; well and good. But does *God* exist? Christopher Hughes has told me that, in his view, if there were a being to which either of these phrases applied, that being would be something very much like the neo-Platonic One[6] – and the One is certainly not God. The One is certainly *not at all like* God.

One might demonstrate the existence of God by supplementing the ontological argument with a proof that that a "something than which no greater can be conceived" or a "supremely perfect being" would have all the "divine attributes" (whatever attributes a definitive list of the divine attributes might comprise, it would certainly contain personality, an attribute that the One conspicuously lacks). Any it is very doubtful whether such a proof is possible: any argument for the conclusion that a supremely perfect being would have all the divine attributes, I think, would have premises that essentially assumed the point at issue – 'Personality is a perfection', for example.

In any case, every known version of the ontological argument is irremediably defective – although the various versions of the argument are defective for different reasons.[7]

The question, "Why do you consider yourself an analytical philosopher instead of a Thomist?" suggests that there is some sort of incompatibility between being an analytical philosopher and being a Thomist. I find this a very puzzling suggestion, for Thomas was himself an analytical philosopher – as were, e.g., Democritus and Aristotle and Descartes. (Consider Thomas and Carnap. Although their philosophical *positions* could hardly have been more different, Thomas's way of thinking about a philosophical question and Carnap's way of thinking about a philosophical question are *far* more similar to each other than either is to any way of thinking that is displayed in the writings of Plotinus, Meister Eckhart, Pico della Mirandola, Fichte, Hegel, Bergson, Santayana, Croce, Ortega y Gasset, Heidegger, Sartre, or Derrida).

There are, of course, current "schools" of philosophy, one of them represented by what goes on in classes and seminars and lectures at the Gregorian University, the University of Louvain, the Institut Catholique de Paris, and the

6 The phrase 'if there were a being …' suggests a difficulty with Hughes's conviction: if there were a being to which these phrases applied, it would, of course be, and neo-Platonists hold that the One is somehow "beyond being".

7 See my essays, "Some Remarks on the Modal Ontological Argument," in Matthias Lutz-Bachman and Thomas M. Schmidt (eds.) *Metaphysik heute*: *Probleme und Perspektiven der Ontologie* (Freiburg and Munich: Verlag Karl Alber, 2007), pp. 132–145 and "Three Versions of the Ontological Argument" in Mirosław Szatkowski (ed), *Ontological Proofs Today*, (Heusenstamm: Ontos Verlag, 2012, pp. 143–162.

Pontifical Institute of Mediaeval Studies – and another by what goes on in classes and seminars and lectures at Oxford University, Princeton University, New York University, and Rutgers University. But the "gap" between these "schools" is more a result of history and tradition (and perhaps of the explicit theological – or atheological – commitments of their members) than of any disagreement about what philosophy is or how it should be "done." And there have been attempts to bridge the gap between the Thomism(s) of the present day and analytical philosophy in the narrow sense (the analytical philosophy of Princeton and Oxford). There is, after all, such a thing as "analytical Thomism" – a term that was, I believe, coined by John Haldane. I am not in a position to provide a list of philosophers who "self-identify" as analytical Thomists, but I think it is plausible to describe the following philosophers as analytical Thomists: Haldane himself, Norman Kretzmann, Cyrille Michon, Roger Puivet, John O'Callaghan, Eleonore Stump, and Mark Murphy.

If I am not a Thomist, this is not because I am an analytical philosopher but because I am a Platonist (or at any rate a "platonist"). Whatever other positions they may hold, Thomists must hold that a human being has as a constituent a substantial form. And the substantial form of Socrates, could not possibly be, or have been, the form of anyone or anything else: anything of which it is the substantial form *must* be Socrates. Consider, for example, Hilary Putnam's thought-experiments involving "Twin Earth". Suppose Twin Earth actually exists. (We are not to include Putnam's stipulation that where there is H_2O on Earth, there is at the corresponding place on Twin Earth a superficially similar stuff with the fictional chemical formula 'XYZ' in our description of Twin Earth: water on *our* Twin Earth is just water.) We have, then, Socrates on Earth and Twin-Socrates on Twin Earth. Thomists maintain that the substantial form of Socrates and the substantial form of Twin Socrates are numerically distinct … well, let's say, numerically distinct *items*. Platonists like myself, however, deny that there is anything that could be called a form of Socrates that is not common to him and to Twin Socrates. There are no doubt many ways to understand expressions like 'the form of Socrates' and 'the form of Twin Socrates', but, however we understand them, whatever precise meaning we give to the operator 'the form of the substance *x*', these two expressions must denote the same thing if they denote anything – a transcendent universal, a universal *ante res*, an item that is in no sense a constituent of either of the two men, a single form in which they both "participate."

In the same vein, a Platonist cannot countenance "individual accidents." Suppose that at mass I observe the celebrant holding a piece of bread. When I look at the bread, Thomas tells us, I do not see the bread – not really, not literally. That is, I do not see the "substance" of the bread. I see only the accidents of the bread: if I see the bread, I see it only vicariously; I see it by seeing (without quali-

fication) its accidents. The priest speaks the words (or some vernacular equivalent of the words), "*Hic est corpus meum*" (with, of course, the intention of doing what priests of the Church have always done in speaking these words at that point in the mass) and the bread – the substance of the bread – is no longer "there"; what is now there is not bread but flesh. The accidents of the bread, however, remain. I see no change in what is in the priest's hand simply because I see there what I saw before those words were spoken, the accidents of the bread. And what is true of sight is true of the other senses: "taste and touch and vision to discern thee fail".

Although no one is a more zealous adherent of the doctrine of the Real Presence than I, I cannot accept Thomas's explanation of what occurs at the moment of Consecration. And my reason for being unable to accept it is not theological but metaphysical – or perhaps one might even say semantical: I can attach no sense to the words 'individual accident'. I am sometimes asked by my Roman Catholic colleagues at the University of Notre Dame (they know I am an Anglican) whether I believe in "transubstantiation". I ask my interlocutors what they mean by the word. They often respond by saying something like, "We mean by it what Aquinas meant by it." And I reply, "I don't even understand what he says about the unconsecrated bread."

9

Mirosław Szatkowski:
Are there any other questions you would like to answer?

Peter van Inwagen:
Oh, eight is a very nice number. If eight Beatitudes sufficed for our Lord, eight questions will suffice for me.

Elisabeth M. Bolduc

Jonathan and Peter

I remember the morning Peter came downstairs noticeably shaking — the phone in his outstretched hand, he was scarcely intelligible. Susan Lowe was calling from Durham to say that Jonathan had died suddenly the previous evening. I took the phone from Peter, and talked with Susan. We ended the conversation with the promise to be in close touch. Peter and I were shattered.

Jonathan's death ended a special friendship between two men who were very different philosophically, but who shared certain values and characteristics. Their wives continue to be friends: I am not a philosopher, but with the help of Susan, I think I can piece together what made Peter and Jonathan care for each other the way they did.

Jonathan (and Susan) and Peter began their friendship in 1994, while the three — having attended a philosophy conference in Salzburg — chatted during a long layover in a lounge in Amsterdam airport. That was when the two came to know each other as people, and not just as philosophers; Jonathan and Peter had met before, but only briefly and professionally. After the Amsterdam chat, subsequent philosophical meetings led to discussions of family and other issues, as well as visits in Durham by Peter, and in South Bend by Jonathan and Susan. When Peter was awarded an honorary degree at St. Andrews, the Lowes were present. Their wives grew closer, as did their blonde daughters. Jonathan's illness in 2013 was worrisome, and Peter arranged to visit Durham and Jonathan that fall, while giving papers in the UK. Tragically and sadly, Jonathan's death came in January of 2014.

What bound these two men in such a special friendship? First and foremost, Jonathan and Peter respected the love and dedication each other had for philosophy, and specifically metaphysics. In Susan's and my view, they were linked through being men of great philosophical integrity. Although they were also both kind people, when the quality of work was involved, they could not — and I mean COULD not — compromise. Further was their shared championing of those who did good work, and their giving of time to philosophers from countries where academic freedom was limited. They knew their encouragement was important; being in the profession, in their view, required that.

Grounding their shared integrity were other qualities that linked these introverted metaphysicians: wide-ranging interests, above average intelligence, and a high value on relationships — especially family. Jonathan and Peter could talk about almost any topic — literature, science — and enjoy themselves immensely.

https://doi.org/10.1515/9783110664812-003

That they also both loved their families and friends deeply was a quality they valued in each other.

I miss Jonathan, and Peter misses Jonathan, and we are joyful to count Susan among our dearest friends. Peter knows — and Susan and I know — that real friendship is most precious, and something to be valued, and treasured, and remembered.

Part II: **The Grounds and Ways of Metaphysics**

Gila Sher

Where Are You Going, Metaphysics, and How Are You Getting There? – Grounding Theory as a Case Study

Abstract: The viability of metaphysics as a field of knowledge has been challenged time and again. But in spite of the continuing tendency to dismiss metaphysics, there has been considerable progress in this field in the 20th- and 21st-centuries. One of the newest – though, in a sense, also oldest – frontiers of metaphysics is the *grounding* project. In this paper I raise a methodological challenge to the new grounding project and propose a constructive solution. Both the challenge and its solution apply to metaphysics in general, but grounding theory puts the challenge in an especially sharp focus. The solution consists of a new methodology, *holistic grounding* or *holistic metaphysics*. This methodology is modeled after a recent epistemic methodology, *foundational holism*, that enables us to pursue the foundational project of epistemology without being hampered by the problems associated with foundationalism.

The viability of metaphysics as a field of knowledge has been challenged time and again. Some have challenged "traditional" metaphysics, or what was considered to be "traditional metaphysics" at the time; others have challenged metaphysics in general. Kant falls under the former category, Carnap under the latter. Kant likened Plato's metaphysics to a "light dove" who, "cleaving the air in her free flight, and feeling its resistance, might imagine that her[1] flight would be still easier in empty space". "Plato", Kant continues, "ventured out beyond [the world of the senses] on the wings of the ideas, in the empty space of the pure understanding. He did not observe that with all his efforts he made no advance – meeting no resistance that might, as it were, serve as a support upon which he could take a stand, to which he could apply his powers, and so set his understanding in motion." (Kant (1781/87), A5/B8-9). Carnap rejected metaphysics altogether: "the so-called statements of metaphysics are *meaningless*"; "metaphysics in its entirety consists of … pseudo-statements" (Carnap (1932), p. 61).

1 I follow Kant's original text by using "her" ("ihr") for the dove in this place, where Kemp Smith uses "its", although in the first part of the sentence he uses "her". Guyer and Wood use "it" in both places. Kant's original formulation of the sentence is: "Die lechte Taube, indem sie im freien Fluge die Luft teilt, deren Widerstand sie fühlt, könnte die Vorstellung fassen, daß es ihr im luftleeren Raum noch viel besser gelingen werde."

https://doi.org/10.1515/9783110664812-004

In spite of the continuing tendency to dismiss metaphysics, there has been considerable progress in this field in the 20th- and 21st-centuries. Both continental and analytic philosophers contributed to this progress, the latter including Prior, Barcan Marcus, Kripke, Lewis, Plantinga, Armstrong, van Inwagen, Stalnaker, Williamson, and others. One of the newest – though, in a sense, also oldest – frontiers of metaphysics is the *grounding* project. Traced back to Aristotle, the grounding project has been recently renewed by Fine (2001, 2012a,b), Rosen (2010), Schaffer (2009), Sider (2011), and others.

In this paper I will raise a methodological challenge to grounding theory – the theory (or theories) developed in pursuit of the new grounding project – and propose a constructive solution. Both the challenge and its solution apply to metaphysics in general, but grounding theory puts the challenge in an especially sharp focus.

1 The grounding project/theory

The *grounding project* is a metaphysical project that seeks to provide an explanatory account of reality in terms of what is grounded in, or depends on, what. This project is often combined with the *fundamentality* project, which endeavors to ground reality in a layer of fundamental elements. In this paper I mean by "the grounding project/theory" the grounding-and-fundamentality project/theory.[2] The grounding project is a vibrant theoretical project, going against the current deflationist, quietist, and philosophy-made-easy trends. It is a "substantivist" project, in the intuitive, common-sense meaning of the word.[3]

The origins of the grounding project, as we have noted above, go back to Aristotle, and in particular to his conception of metaphysics as providing an explanatory description of reality based on the idea of ontological priority. Thus, in introducing his work on grounding, Schaffer says:

> I will argue for the revival of a … traditional Aristotelian view, on which metaphysics is about what grounds what. (Schaffer (2009), p. 347)

2 In speaking about contemporary works on metaphysical grounding I will alternate between "theory" and "project", depending on which perspective on this work I wish to emphasize.
One grounding theorist who does not require fundamental elements is Rosen (2010). Since much of what he says, however, falls under my category of grounding theory/project, I will include his work in this category.

3 Compare with the substantivist approach to truth (e.g., Sher (1998, 2004, 2016b)).

And describing what Aristotle's view amounts to, he says:

> [O]n Aristotle's view, metaphysics is the discipline that studies substances and their modes and kinds, by studying the fundamental entities and what depends on them. (*Ibid*., p. 351)[4]

While the different practitioners of the grounding project differ on various points, several characteristics emerge as central to this project, as it is currently pursued. Four of these are:[5]

1. *The ideas of dependence and fundamentality are central to grounding*. The idea of substantive dependence is the idea of what depends on what. It is the main idea underlying the typical vocabulary of grounding: "in virtue of", "because", "explains why", "is due to the fact that", and so on. The idea of fundamentality is the idea of what is basic, namely, what the ultimate elements of the dependence relation are.

Ground theorists emphasize the centrality of dependence and fundamentality both to grounding theory and to metaphysics more generally. Thus, the title of Rosen's paper on grounding is "Metaphysical Dependence: Grounding and Reduction" (Rosen (2010), p. 109). And a subsection in Schaffer's article on grounding is titled "*Ordering: The Importance of Dependence Structure*" (Schaffer (2009), p. 362). Schaffer sums up his paper by saying:

> [M]etaphysics as I understand it is about what grounds what. It is about the structure of the world. It is about what is fundamental, and what derives from it. (*Ibid*., p. 379)

And Sider says:

> Metaphysics, at bottom, is about the fundamental structure of reality. ... the ultimate goal is insight into ... what the world is like, at the most fundamental level. (Sider (2011), p. 1)

2. *Grounding is strongly hierarchical*. The grounding relation, *X grounds Y*, is strongly hierarchical. In this paper, I will understand by a "strongly hierarchical grounding relation" a partially-ordered grounding relation – anti-reflexive, anti-symmetric, and transitive – with minimal ("fundamental") elements, where each non-minimal element is grounded in minimal element(s) in a finite number

4 Fine (2012b), p. 8, fn. 1) also indicates that his "conception of metaphysics is broadly Aristotelian in character".

5 Note: These characteristics hold regardless of whether we identify the units of grounding and fundamentality as facts, propositions, truths, entities (objects), etc. Since my own concerns in the present paper are also independent of this question, I will put it aside here.

of steps. I will call such a relation a "strictly-ordered" or a "strictly hierarchical" relation.

Although grounding theorists differ in the extent to which they offer a detailed description of the formal structure of the grounding relation as well as the specific features they attribute to it, they all view it as strongly hierarchical:

> [T]he attempt to determine what grounds what naturally proceeds in stages – one first determines the relatively immediate grounds for the truths in question, then the relatively immediate grounds of those grounds, and so on until one reaches the ultimate grounds. (Fine (2012a), p. 44)

> [T]he relation [of ground is] irreflexive and anti-symmetric. (*Ibid*., p. 45)

> [The contemporary philosopher of grounding] will begin from *a hierarchical view of reality* ordered by *priority in nature*. The primary entities form the sparse structure of being, while the grounding relations generate an abundant superstructure of posterior entities. (Schaffer (2009), p. 351)

> Grounding is … irreflexive, asymmetric, and transitive. It thus induces a partial ordering over the entities (*the great chain of being*) … . Formally this may be modeled by a directed acyclic graph, for which every path has a starting point. (*Ibid*., p. 376)

> [T]he fundamental facts underwrite or give rise to all other facts. (Sider (2011), p. 105)

And Rosen says that "the binary part of the grounding relation is asymmetric and hence irreflexive". He then characterizes these features as "[s]trong asymmetry" and "[s]trong irreflexivity". He also assumes "transitivity in a strong form". He indicates that "the [grounding] relation is presumably not connected", so we have only a "partial order" (Rosen (2010), pp. 115–116). In one of his examples – that of a naturalistic grounding – he identifies the grounding relation with a (mathematical) tree: "every fact tops a naturalistic tree" (*Ibid*., p. 112).

In an encyclopedia article on metaphysical grounding, Bliss and Trogdon describe the grounding relation as "*well-founded*" (Bliss and Trogdon (2014), p. 10).

3. *Grounding is objective*. What grounds what and what is basic or fundamental are objective matters, not just in the sense of being intersubjective, but also, and most importantly, in the sense of being *factual*, that is, being features of the world (reality) itself. Fine speaks about ground as a relation between *worldly entities* such as *facts* (Fine (2012a)), and he emphasizes the connection between grounding and realism (Fine (2001)). Rosen (2010), too, regards the grounding relation as a worldly relation among facts. Sider (2011) titles his book on grounding "Writing the Book of the World" and continuously emphasizes his interest in the *objective structure of the world*. There is a "fundamental structure of reality" (*Ibid*., p. 1),

Sider says, and the grounding or fundamentality project is that of uncovering this structure. Schaffer, too, as a neo-Aristotelian, talks of grounding in terms of substances, objects, entities, existence, and world: "metaphysics as I understand it is about what grounds what. It is about the structure of the world" (Schaffer (2009), p. 379).

Another aspect of the objectivity of grounding is veridicality or truth. The grounding project, as I understand it, is subject to a robust veridicality requirement. Grounding theory is required to provide a *true* description of what grounds what, where "true" is understood in a strong sense, closer to correspondence than to coherence or pragmatic truth.

4. *Grounding theory is highly explanatory*. The task of grounding theory is to provide a substantial and highly explanatory account of reality in terms of grounding. This point is salient for all the grounding theorists we are considering:

> We take *ground* to be an explanatory relation: if the truth that *P* is grounded in other truths, then they *account* for its truth. (Fine (2001), p. 15)
>
> [T]he relationship of ground is a form of explanation; in providing the ground for a given proposition, one is explaining, in the most metaphysically satisfying manner, what it is that makes it true. (*Ibid*., p. 22)
>
> [Ground is] a distinctive kind of metaphysical explanation, in which explanans and explanandum are connected … through some constitutive form of determination. (Fine (2012a), p. 37)
>
> [T]he grounding relation is an explanatory relation – to specify the grounds for [*p*] is to say why [*p*] obtains. (Rosen (2010), p. 117)[6]

Schaffer (2009) contrasts his Aristotelian conception of ontology with Quine's conception which limits ontology to a mere "*list*" of "*beings*" (*ibid*., p. 348). And Sider (2011) views his entire metaphysical project (with its notions of *structure, carving at the joints, fundamentality*, and *grounding*) as substantive and explanatory:

> [The book] show[s] how structure illuminates explanation, …, substantivity, … . (*Ibid*., p. ix)
>
> Good … theories … must be cast in joint-carving terms. We may put this in terms of explanation: "theories" based on … non-joint-carving classifications are unexplanatory. (*Ibid*., p. 23)[7]

6 Here "[*p*]" stands for "the fact that *p*".

7 This citation is explicitly about good scientific theories. But it is quite clear that it holds for all

> Fineans and I can give satisfying ultimate explanations. For we accept structured and plentiful fundamental truths, and can tell detailed stories about how they ground (Fine) or are metaphysically truth-conditions for ([Sider]) various nonfundamental truths. (*Ibid.*, p. 161)

2 The grounding project and the foundationalist project

Although the grounding project is a metaphysical, largely descriptive project whereas the foundationalist project is an epistemic, largely justificatory project, it is hard not to see significant similarities between the two.[8]

The foundationalist project is a well-known epistemic project, so there is no need to describe it here in detail. A classical example of this project is Descartes's *cogito* project. A later example is Frege's and Russell's logicist project, and more recently, we may view some forms of naturalism (see example below) as foundationalist in character.

Briefly, the foundationalist project is a theoretical philosophical project that seeks to construct an objective and highly explanatory foundation for human knowledge. Human knowledge, here, includes the totality of our theories of the world (various facets of the world), or, on a more mundane level, our claims about the world.[9]

The foundationalist project shares the four distinctive characteristics of the grounding project described above: the centrality of *dependence*, the requirement of a *strongly hierarchical structure*, the demand of *objectivity*, and the commitment to a *highly explanatory* account. These features can, in principle, characterize both descriptive and justificatory projects, both metaphysical and epistemic projects. And they do characterize both the grounding and the foundationalist projects. We have seen how they characterize the grounding project. Their characterization of the foundationalist project is straightforward:

1. The relation *X founds Y* is a dependence relation: If *X* founds *Y*, then *Y* (or *Y* having the status of knowledge, or the justification of *Y*) depends on *X*.

---

theories, including metaphysical theories, according to Sider.

8 A similar point is made by Thompson (2014).

9 (i) The understanding of "world" may vary from one foundationalist to another, and such variations have significant consequences for the proposed foundations, but for the most part, the general principles remain the same.
(ii) For accounts of foundationalism that are similar in spirit to the one given below, see, e.g., Sosa (1980a,b).

2. This relation is also an objective relation: If *X* founds *Y*, then *X* is real or objective and it objectively founds/justifies *Y*.
3. The founding relation is (or is required to be) strictly hierarchical as well: The relation *X founds Y* is a strong partial-ordering. It is anti-reflexive, anti-symmetric, and transitive. It has minimal (basic, foundational) elements, and each non-minimal element is connected by finite chains to the minimal elements that form its ultimate foundation.
4. Finally, the founding relation is required to be highly-explanatory: If *X* founds *Y*, then *X*, along with its (founding) relation to *Y*, explain how *Y* is justified (or why *Y* is a genuine item of knowledge).

The foundationalist project has other characteristic features as well. For example, it requires absolute certainty of the founding of knowledge. But since this feature is not shared by the grounding project, it is of lesser interest for us here. One result of the strict-hierarchy requirement of the foundationalist project is that it bans all forms of circularity and infinite regress.

We see that, their differences notwithstanding, the foundationalist and grounding projects are similar in several significant respects. In particular, they both share the four characteristics noted above: both projects aim at being, and claim to be, objective and highly explanatory, and their central relations, *X grounds Y* and *X founds Y*, are both strongly hierarchical dependence relations. Occasionally, the grounding and foundation relations extensionally coincide. An example of a metaphysical grounding-chain (due to Fine (2012a), p. 44) that is also an example of a foundationalist (epistemic) grounding-chain is:

> The Normative *is grounded in* the Natural;
> The Natural *is grounded in* the (Macro-) Physical;
> The (Macro-) Physical *is grounded in* the Micro-physical.

The foundationalist project, however, is fraught with difficulties, and today many epistemologists regard this project as flawed beyond repair. Here I will focus on one of its serious problems, which has to do with its shared characteristics with the grounding project.

The foundationalist project requires that the founding or justification relation be strictly hierarchical. But if the justification relation is strictly hierarchical, then the main burden of justification falls on the minimal elements of this relation, namely, on the *founding* elements of the foundationalist hierarchy. If the minimal (founding) elements lack appropriate justification (foundation), then our entire body of knowledge lacks justification (foundation). What is an appropriate justification? Two central requirements on an appropriate justification are, as we

have seen above, objectivity and explanatory power. So the main burden of an objective, explanatory justification (foundation) falls on the minimal, foundational elements. But where could the minimal elements get their objective, explanatory justification from? There are no elements lower than the minimal elements in the foundationalist hierarchy, hence there are no elements that could provide, or could produce resources for providing, an objective, explanatory justification of the fundamental elements.

Foundationalists may say that foundational elements do not require an objective, explanatory foundation or justification. Explanation and justification must stop somewhere; it is impossible to either objectively justify or give an explanatory account of everything, and foundationalism is not to be blamed for not doing the impossible. But this claim is problematic in several respects. For one thing, the question is not whether the foundationalist project can objectively and explanatorily justify everything, but whether it can establish and explain something very specific, namely the foundational elements that play an active role in founding the rest of our knowledge. Not all elements are alike. Failure to give an objective, explanatory justification of an isolated unit of knowledge will not undermine the entire foundationalist project, but failure to justify foundational elements that are supposed to found, directly or indirectly, many nonfoundational elements, will. A system of knowledge grounded in unfounded elements is like a building having a "foundation" of sand. It is important to note that this problem is independent of the absolute certainty requirement of the foundationalist project. Even if we do not require that the foundational elements be founded in a perfect, complete, once-and-for-all way, with no possibility of error, the problem remains. It remains even if all we require is significant progress toward an establishing and explaining the foundational elements.

Furthermore, just because it is impossible to do *Y* does not mean that a project *X* cannot be criticized based on its inability to do *Y*. If *X requires* something that is in principle undoable, this is a reason to question *X*, not to excuse it. It is a sign that something is wrong with *X*. If the viability of the foundationalist project requires that the foundational elements be objectively and explanatorily justified, then if, in principle, this requirement cannot be satisfied, this casts doubt on the foundationalist project. A project that, to be viable, must do the impossible, is not viable.

The foundationalist may concede these points but say that the justification of foundational units of knowledge is inherently different from that of the other elements: they are, in principle, justified without using any other units of knowledge. They are self-justifying, or else they are justified without resort to any knowledge whatsoever. Self-justification violates one central principle of foundationalism: its ban on circularity. The logical prototype of a self-justifying item of knowledge

is: "Φ; therefore Φ". While such justification is objective in the sense that it is factually (or, indeed, logically) valid, it fails both to establish the objectivity of Φ and to explain its ability to found other items of knowledge. As far as the validity of "Φ; therefore Φ" is concerned, Φ might be a figment of our imagination and its ability to justify other units of knowledge might be null.

Four contenders for a "self-standing" justification, i.e., justification that does not appeal to any unit of knowledge, are: pure sensory perception, intuition (either everyday intuition or rational intuition), common-sense obviousness, and conventionality. But all four are highly problematic both with respect to their objectivity and with respect to their explanatory power. The epistemic credentials of pure sensory perception were criticized by, e.g., Sellars (1956), under the heading "the myth of the given". The epistemic credentials of intuition were questioned by e.g., Benacerraf (1973), Harman (1977), and Cummins (1998). Those of common-sense obviousness were criticized by, e.g., Sher (1999). And Quine (1935, 1954) sharply criticized conventionality as trivializing the very idea of knowledge. For additional criticisms of all these contenders, see Sher (2016a), Chapters 2 and 9.

To give the flavor of some of these criticisms, take common-sense obviousness as an example. A few (not necessarily disjoint) criticisms of common-sense obviousness as a source of foundational knowledge are: (a) Our sense of obviousness is often utterly unreliable. (Think of what was considered obvious prior to the revolutionary discoveries of modern science and mathematics.) (b) It is not clear in what way common-sense obviousness is said to justify the foundational elements of knowledge. If the claim is that all obvious elements are foundational (minimal), it is false. If it is that all foundational elements are obvious, it requires objective justification and explanation. (c) The foundationalist project is a theoretical rather than a phenomenological or a psychological project; hence the justification of the foundational elements has to be theoretical. A justification based on common-sense obviousness, however, does not satisfy this requirement. It is impressionistic or psychological, but not theoretical. (d) Obviousness is an exceedingly weak, unobjective, and unexplanatory standard of fundamentality. In short, common-sense obviousness cannot do any of the things that an objective and explanatory theoretical foundation must do.

In light of the similarities between the grounding project and the foundationalist project, the question arises whether the former suffers from some of the problems that undermined the latter. These similarities, as we have seen above, center on four central characteristics of the grounding project – *centrality of dependence*, *objectivity*, *explanatory power*, and *strict hierarchy* – and the question is whether the strictly hierarchical structure of the grounding project subverts its goal of a highly-explanatory, objective account of reality in terms of dependence. Unfortunately, the answer to this question appears to be positive. If the grounding-

account of reality is strictly hierarchical, then the main burden of its objectivity and explanatory power falls on the minimal, fundamental elements. If the fundamental elements are deprived of objectivity and resist explanation, then the entire grounding falls short of objectivity and explanation. Suppose the fundamental elements are arbitrary, suppose they are figments of our imagination, suppose they are irrevocably mysterious, or their ability to ground other elements is magical. In all these cases the grounding of higher elements will ultimately lack both objectivity and explanatory power. Suppose *X* is grounded in a fundamental element *Z* through an intermediate element *Y*. Without establishing the objectivity of *Z*, without understanding what *Z* is (what its features, laws, and/or principles or regularities are), without establishing that *Z* in fact grounds *Y*, and without explaining how it grounds *Y*, the grounding of *X* has very little objectivity and explanatory power.[10]

How could grounding theorists handle this problem? Responses analogous to those attempted by foundationalists – saying that we have no choice but to leave some elements unestablished/unexplained, appealing to common-sense obviousness, sensory perception, intuition, or conventionality – will not do here too, and for similar reasons to those given in the case of foundationalism. (Although here the crux of the matter is theoretical description rather than theoretical justification, the requirements of objectivity and explanatory power will be violated here too.)

In the next section I will propose an adjustment to the grounding project that will solve the problem without undermining the project itself.[11] This solution is analogous to one I recently proposed in response to the above-mentioned problems with the foundationalist project. In the case of grounding, however, the adjustment can preserve more features of the original project than in the case of the foundationalist project.

10 The inadequacy of having unexplainable fundamental elements is also noted by Chang (2013), though her point is specific to a particular context of grounding: the grounding of practical reasons. When we reach the fundamental elements of the grounding, Chang says, "there's no more explanation to be had, end of story. Facts that are explanatorily primitive are self-grounded; they cannot be accounted for in any other terms and represent the end of the line in explanation." (*Ibid.*, p. 165) The problem with self-grounded facts, according to Chang, is "the *Problem of Explanatory Shortfall*" (*ibid.*, p. 170). In some cases "it is wholly unsatisfying to rest with 'That's just how things are'." (*Ibid.*, p. 173)

11 Needless to say, I do not claim this is the only possible solution to the problem.

3 Holistic grounding

My solution to the problem facing grounding theory is methodological. I will propose a new methodology, *holistic grounding*, that makes an objective and highly explanatory account of the fundamental elements possible. Holistic grounding is modeled after *foundational holism*, an epistemic methodology developed in Sher (2016a) and designed to avoid the foundationalist predicament.[12] Both foundational holism and holistic grounding involve a special kind of holism, one that differs from most existent conceptions of holism. I will begin with foundational holism and then turn to holistic grounding.

3.1 Foundational holism

The key to understanding foundational holism as an alternative to foundationalism lies in distinguishing between the concepts of *foundation* and *foundationalism*. A *foundation for knowledge*, under this distinction, seeks to establish the viability of human knowledge, both empirical and abstract, and provide objective, veridical, and highly explanatory justification of such knowledge. A *robust* foundation establishes our claims to knowledge by connecting them to the *world*, thus by exhibiting an ultimate dependence of our knowledge on the world. Three of the four characteristics we have examined in this paper are thus built into the idea of a *robust foundation*: (i) dependence, (ii) objectivity, and (iii) strong explanatory power. But the fourth characteristic – a strictly hierarchical justification relation – is not part of the idea of a foundation. This characteristic has to do with the *methodology* used to pursue the foundation project, and in principle different methodologies might be used in pursuit of this project. Foundationalism and foundational holism are two distinct methodologies for pursuit of the foundational project.

One of the distinctive characteristics of the foundational*ist* methodology is its requirement that the foundation of knowledge be strictly hierarchical. Foundational holism renounces this requirement. Another distinctive feature of the

12 There are some similarities between foundational holism and foundherentism (Haack (1993)), but there are also significant differences between them. Two of these are: (a) While foundherentism is limited to empirical knowledge, foundational holism is applicable both to empirical and to abstract (e.g., logical and mathematical) knowledge. (b) Foundational holism is *holistic* rather than *coherentist*. (The *holism in foundational holism* is, as we shall see below, not a coherentist holism.)

foundationalist methodology is its requirement that the foundation be absolutely certain. This requirement, too, is renounced by foundational holism.

Viewing foundational holism as a project, namely the project of *foundation-without-foundationalism*, two of its main principles are:

(a) Every field/item of knowledge, qua a field/item of knowledge, requires a robust, objective, and highly explanatory foundation in the world (broadly understood) or in those facets of the world that it targets.
(b) The founding/justification relation may take different forms in different cases and at different times. The underlying idea is that the foundational project is a dynamic project. There are multiple ways for our theories to reach, and be founded in, the world, some simple, others complex, some strictly hierarchical, others not. What pattern the justification relation can/should take is affected by particular circumstances, including the "distance" between the targeted facets of the world and our cognitive resources for reaching these facets. The point is that some facets of the world are more difficult for us to discover than others, given our cognitive resources, and some theories are more difficult to justify, requiring more complex (circuitous, indirect) patterns of justification than other theories.

These principles point to two ways in which foundational holism differs from other conceptions of holism. First, it is world-oriented rather than coherentist. While coherentist holism says that the justification of an item of knowledge largely consists in establishing its coherence with other items of knowledge, foundational holism says that it primarily consists in establishing its connection to the world. Second, foundational holism licenses the use of *two* rich networks of interconnections by the foundational project: (i) a network of connections among fields (theories, items) of knowledge, and (ii) a network of connections between fields (theories, items) of knowledge and the world. The two networks themselves are interconnected. Most importantly, the first network enriches the second and its interconnections are integrated into those between our body of knowledge and the world.

But foundational holism differs from other conceptions of holism in other ways as well. For example, one conception of holism regards it as "wholistic" in character. Dummett (1973/81) calls this type of holism "total" holism and I call it "one-unit" holism (Sher (2016a)). One-unit holism is the view that the smallest unit of knowledge is our body of knowledge as a whole.[13] In contrast to this type of holism, foundational holism regards our body of knowledge as consisting of

13 Dummett (1973/81) and Glymour (1980) attribute this type of holism to Quine.

multiple elements. It is a *network of independent elements*, *standing in multiple relations*. Another conception of holism regards it as "unstructured". Unstructured holism is the view that every item of knowledge is equally connected to any other item of knowledge.[14] Foundational holism, in contrast, is a *structured* holism. It says that the (epistemically relevant) connections between items of knowledge are as selective, highly structured, and systematic as they are openended.

One feature that foundational holism shares with other types of holism is its attitude toward circularity: it does not ban all forms of circularity. More specifically, foundational holism distinguishes four types of circularity: *destructive*, *trivializing, indifferent*, and *constructive*. Destructive circularity is the type of circularity that leads to paradox. Some cases of self-reference (e.g., "the set of all sets that are not members of themselves") may fall under this category. A paradigmatic example of trivializing circularity is "*P*; therefore *P*". A justification of *P* by a logical inference from the assumption that *P* is trivial to the point of not counting as a justification. These two kinds of circularity are banned by foundational holism as much as by foundationalism. Indifferent circularity is the circularity involved in studying English grammar using English grammar. It is neither better nor worse to study English grammar in a language that uses English grammar than in a language that uses, say, French grammar. Constructive circularity is an instrument of knowledge. Gödel's representation of syntax by syntax, Henkin's syntactic model of standard 1st-order logic, and other achievements in set theory, meta-mathematics, and meta-logic make ingenious use of patterns that have circular elements. Foundational holism regards constructive circularity as an invaluable epistemic tool.

One key to productive uses of circularity is *partiality*. The knowledge obtained is only partially circular. Non-circular elements are also involved and play a significant role. Thus, consider Henkin's use of 1st-order syntax to prove the syntactic (proof-theoretic) completeness of 1st-order logic. Many other elements, including semantic principles and mathematical (set-theoretical) laws play a crucial role. Another key is a *discerning* use of circularity. Consider Russell's discovery of a paradox in Frege's logic. Given that Russell's paradox involves relations and multi-quantifier quantifier-prefixes, he had to use a quite powerful logic to discover the paradox, and at the time the only powerful logic available to him was Frege's logic (or some variant of Frege's logic).[15] But whatever elements of Frege's

14 Friedman (2001) attributes this type of holism to Quine as well.

15 For discussions of how Russell discovered his paradox, see, e.g., Grattan-Guinness (1978), Coffa (1979), and Moore (1988). But these articles do not raise the question of what logic Russell used to discover the paradox.

logic Russell used to discover the paradox, he used them flexibly, dynamically, critically, and intelligently – holding off some elements, switching from some elements to others, and so on – so the paradox could come to light.[16]

One project in which some measure of circularity is unavoidable is the foundational project, and in particular those parts of this project that deal with "basic" elements, elements that significantly contribute to the founding of most other elements.[17] For example, one cannot provide a foundation for logic without using logic. But by heeding the principles of partiality and discerning-use, a foundation for logic is made possible. Thus, the elements that do the major work in the holistic foundation for logic delineated in Sher (2013, 2016a) are not logical. They are philosophical and mathematical, and the work significantly involves general knowledge, nonlogical principles of rationality, all-purpose intellectual activities (such as "figuring out")[18], and so on. The foundation proceeds in a series of questions that are quite independent of the (background) logic used in answering them: "What is the task of logic in our system of knowledge?", "Can logic be grounded only in the mind (language, concepts) or does it require a grounding in the world as well?", "Why does logic require a grounding in the world?", "What specific features of the world are capable of grounding logic, how and why?", "What are the sources of the generality, necessity, and normativity of logic?", "What is the relation between logic and mathematics?", and so on. None of these questions or the answers given to them center on logical claims. The foundation employs elements from a variety of fields of knowledge, including logic, but its structure is as far from "*P*; therefore *P*" as that of any worthwhile scientific, mathematical, or philosophical theory.[19]

It is important to note that although the foundational holistic method renounces the strict-ordering requirement of the foundationalist methodology, it

16 (i) The discoveries of the liar and heterological paradoxes are also arguably of this kind.
(ii) For a similar view of circularity as potentially productive see Sosa (1997).

17 One characteristic of foundational-holism is that the relation *X plays a significant role in founding Y* is not transitive. *Z* may play a significant role in founding *X*, but once we get to *Y* so many other elements might be involved in founding it that the role of *Z* can cease to be significant, or simply, in the context of *Y*, *Z* is no longer very relevant.

18 I use "figuring out" as a general term for a cluster of activities, from a baby figuring out how to make the mobile on her crib move (e.g., by hitting the bed with her feet (so it shakes)), a technician figuring out why a certain instrument is not operating properly, a mathematician figuring out how to solve a certain mathematical problem, and so on.

19 I should add that the general character of the above questions does not rule out precise results. For example, the answers given to these questions in the above-mentioned works lead to a precise criterion of logicality. We will briefly discuss this criterion below.

neither rejects nor denies the advantages of strictly-ordered founding, or strictly-ordered sections of the (overall) founding process. An example of a strictly-ordered (-hierarchical) justification sanctioned by foundational holism is a logical proof (that is, the series of steps involved in a logical proof). The foundational holistic method embraces strictly-hierarchical justification, but it also says that when such justification comes to an end, this is not the end of theoretical justification. Other patterns of justification are available as well, and these enable us to engage in extensive foundational projects that are rational, objective, highly explanatory, and critical, yet not strictly hierarchical (or not strictly hierarchical through and through).

3.2 Holistic grounding

In light of the inherent similarities between the foundationalist and grounding projects – both their common characteristics and their analogous problems with the minimal elements of the respective hierarchies – it is reasonable to expect that a solution to the minimal-elements problem of one project could be adapted to the second project. I will call an adaptation of foundational holism to the grounding project "holistic grounding". Holistic grounding can be developed in a number of ways. In particular, it can be developed in ways that render it a friendly amendment to the current conception of the grounding project and in ways that render it an alternative to that conception. The crux of the matter is whether holistic grounding preserves the strict-hierarchy requirement for the non-fundamental elements, limiting the holistic treatment to the fundamental elements, or whether it views the grounding of all elements holistically. The holism described above in connection with foundational holism is, as we have seen above, compatible with giving a preferred status to hierarchical grounding whenever this is a viable option, but it is also compatible with giving equal status to hierarchical and nonhierarchical grounding.

Given the importance of *objectivity* and *explanatory power* for the grounding project, it is reasonable to use these requirements as a touchstone in determining the balance of hierarchical and non-hierarchical patterns in holistic grounding. The objectivity gauge is associated with such questions as: "Is the structure of reality in fact strongly hierarchical?", "Is it strongly hierarchical in all areas or just in some areas?" The explanatory-power gauge is associated with questions like: "Is a hierarchical or a non-hierarchical grounding-description of reality more explanatory in case/area X?".

I will not attempt to answer these questions here; the answers to these questions and the precise development of holistic grounding as a metaphysical meth-

odology and descriptive project require an independent paper. Instead, I will briefly report on a few considerations that led other philosophers in the direction of a holistic approach to metaphysics and propose an example of holistic grounding in one fundamental field, logic.

Barnes (forthcoming) points out, or argues for the putative reality of, a few cases of nonhierarchical dependence:

1. Electrons, as universals, depends on their instances, and their instances depend on electrons as universals corresponding to natural kinds. (*Ibid*., p. 9)[20]
2. Armstrongian "[s]tates of affairs depend on – and are thus explained by – their constituents" (particulars and universals) but the reason the constituents exist is that they constitute states of affairs. The "individual constituents depend on – and are thus explained by – states of affairs". Barnes calls this "explanatory holism". (*Ibid*., p. 10)[21]
3. "[T]here are tropes which mutually depend on each other. You cannot have a mass trope without a size trope and a shape trope, for example. … The picture here is one of 'dependence clusters' – mass depends on shape and size, size depends on mass and shape, etc." (*Ibid*., p. 11)[22]
4. On the realist, structuralist conception of numbers as places in a structure, each number depends on the other numbers, since its place in the structure depends on their places. (*Ibid*., p. 12)[23]
5. On an inflationary metaphysics of events, larger events consist of smaller events. Generally, there are larger and smaller events such that the smaller events are essential for the identity of the larger events and vice versa. For example, the (event of the) evacuation of Dunkirk is essential for the identity of (the event of) World War II and (the event of) World War II is essential for the identity of the (event of the) evacuation of Dunkirk. The two are dependent on each other.[24]

The moral Barnes draws from the pervasiveness of such examples is that there is room for *holistic explanation* in terms of dependence. For additional examples of symmetric dependence see Thompson (2014), who uses the term "metaphysical interdependence" (*ibid*., p. 69) for non-anti-symmetric dependence. While Barnes and Thompson agree on holistic dependence, they differ with respect to holistic grounding. Grounding, according to Barnes, is essentially hierarchical; Thompson, in contrast, allows holistic grounding. Another philosopher who introduces some holistic elements into his conception of grounding is Dasgupta

20 Barnes presents this as a neo-Aristotelian conception of dependence.

21 Barnes views this as "the most stable way of making sense of the fact-based ontology that Armstrong wants to defend." (*Ibid*., p. 10)

22 Barnes regards this view, which she traces to Denkel (1996) and Simons (1994), as appropriate for a trope bundle theory.

23 Barnes refers to Linnebo (2008) for this view.

24 Barnes directs us to Hornsby (1997) for this case of symmetric dependence.

(2014). Dasgupta regards grounding as irreducibly plural, where by this he means that it is *clusters of elements*, rather than single elements, that stand in the grounding relation. These clusters of elements are presumably interconnected, hence his view is at least partially holistic.

From the present perspective, however, we are especially interested in the holistic grounding of (what current grounding theorists view as) fundamental elements. Take logic. Sider (2011) considers the logical constants as metaphysically fundamental based on indispensability considerations. But he is unsure where to locate their fundamentality. Should we treat all logical constants on a par with respect to fundamentality, or should we sort them out into fundamental and non-fundamental constants? The question is especially acute in the case of the logical connectives, due to their equal status as truth-functional or Boolean connectives. For Sider, the touchstone of fundamentality is *carving reality at the joints*, where joint carving involves capturing the real, or objective, structure of reality. It is the joint-carving notions that are minimal or fundamental, and a central task of metaphysics is to study the fundamental elements.

The question which logical connectives carve reality at the joints leads Sider to consider several options. One of these is that logical connectives, or logical constants more generally, are non-fundamental, that they are grounded in more fundamental elements, elements for which the above conundrum does not arise. This option, Sider notes, is available in the case of measurement. If we ask: "Which function from pairs of points of space to real numbers is fundamental: the distance-in-meters function, or the distance-in-feet function, or a function corresponding to some other unit?" (*ibid*., p. 217), we have the option of answering: "none of them is; the fundamental metrical facts are facts of spatial congruence" (*ibid*.). But he is skeptical that a similar route is open for logical constants: "Unfortunately, escape of this sort seems unlikely in the case of logic: what more fundamental theory could we shift to?" (*ibid*., p. 218)

It is in cases like these that the power of holistic grounding is most striking. Holistic grounding opens up new possibilities for the grounding of logic. One of these is grounding logic holistically in something more fundamental, from the point of view of carving reality at the joints, than logic itself (viewed as a method or a theory of inference). In Sher (2013, 2016a) I described such a grounding of logic, based, in a holistic spirit, on joint epistemic and metaphysical considerations. These considerations have to do with issues raised by the questions noted in Subsection (3.1) above: "What is the task of logic in our system of knowledge?", "Can logic be grounded only in the mind (language, concepts) or does it require a grounding in the world as well?", "Why does logic require a grounding in the world?", "What specific features of the world are capable of grounding logic, how and why?", and so on.

According to this account, logic in general, and logical constants in particular, are grounded in the formal structure of reality. Logical constants are grounded in formal properties (relations, functions) – the distinguished parameters of formal structure. "Formal structure" is a joint-carving notion in Sider's sense, and "formality" is given a precise, objective, and highly explanatory definition or criterion. This criterion is holistic in the sense that it employs notions, and utilizes knowledge and insights, from various fields.

I will not be able to describe the grounding of logic in the formal structure of reality in detail here (for a detailed account see Sher *op. cit*. and Sher (1991)). But in a nutshell, the idea is that due to the special character of formality (specified by its criterion), formal structures are governed by especially strong laws. If, then, the logical structure of sentences represents the formal skeleton (structure) of the situations they correspond to, and if logical rules of inference represent laws governing formal structures, then logical inferences will be grounded in formal laws governing the formal structures of the situations corresponding to their premises and conclusion. Logical constants, on this account, represent formal properties (relations, functions) of objects (actual or counterfactual), and the criterion of formality (formal properties) ensures that formal laws are sufficiently powerful to ground logic, given its task.

The criterion of formality is an *invariance* criterion. Invariance criteria (sometimes referred to as "symmetries") are highly informative and play a central role in mathematics and science. In the present context we talk about invariance of *properties*. Every property has some degree of invariance, but properties differ in their degree of invariance. The degree of invariance of, say, the property *x is a person* is greater than that of *x is a woman*. *x is a person* is not affected by (does not notice) replacements of women by men, but *x is a woman* does. Formal properties are distinguished by their especially strong degree of invariance. They are *invariant under all isomorphisms* of relevant structures.[25] (In the literature, they are also said to be "invariant under bijections".) In simple terms, the invariance criterion of formality says that a property is formal iff it does not distinguish between isomorphic structures of objects of appropriate types. For example, the identify relation does not distinguish between isomorphic structures of the type $< D, < a, b >>$, where a,

25 A structure is a pair, $< D, \beta >$ where D is a (non-empty) domain of objects and β is an element or an n-tuple of elements of/on D – objects in D or extensions of properties and relations (of any level) in D. (Properties of level 1 are properties of objects, properties of level 2 are properties of properties/relations of level 1, and so on.) Two structures, $S_1 =< D_1, \beta_1 >$ and $S_2 =< D_2, \beta_2 >$, are isomorphic iff (if and only if) one is isomorphic to the other. S_1 is isomorphic to S_2 iff there is a 1-1 and onto function (bijection) f from D_1 to D_2 such that β_2 is the image of β_1 under f.

b are objects in the domain *D*, and as such it is a formal relation. Logical constants denote (stand for, represent, correspond to) formal properties, and any formal property is an admissible denotation of a logical constant.

Formality in the invariance sense explains the logicality of the existential and universal quantifiers as well. The properties corresponding to these quantifiers are the 2nd-level properties of non-emptiness and universality (or universality in a given domain), and these properties satisfy the invariance criterion of formality. Formality (in the above sense) also explains the logicality of the logical connectives in the context of open formulas. (E.g., it explains the logicality of "&" in the context "*Px* &*Qx*" by the formality of its denotation, the intersection operation, ∩.) In the context of sentential logic, where the smallest units are atomic sentences, the formality criterion is generalized. Roughly, connectives are formal iff they are invariant under 1-1 replacements of atomic situations (facts, states of affairs) that preserve the feature of being the case. This criterion of formality coincides with truth-functionality.[26]

We can now say that metaphysically, the notion of logical constant in general is grounded in the notion of formality just as the notion of unit of measure is grounded in the notion of congruence. The notion of formality is an objective notion, and it is given a highly explanatory account in terms of invariance.

The explanation of formality is made possible by our holistic methodology. We use mathematical notions, which are (directly or indirectly) formal, to formulate the invariance criterion of formality. But the circularity in question is constructive. We explain why logicality is grounded in formality and why the invariance criterion is an adequate criterion of formality in terms of a cluster of notions, many of which are not formal.

The notion of formality is highly explanatory in two directions: (a) it is given a highly explanatory account in terms of invariance and (b) it provides, or partakes in providing, a highly explanatory account of other notions, for example, the notion of logical constant. But that is not all. The formality of logic enables us to explain its distinctive characteristics beyond its logical constants: its strong necessity, generality, normativity, apriority (or, in my preferred view, quasi-apriority). Formality plays a central role in the grounding of mathematics as well, leading to a new, highly explanatory account of its interrelations with logic. (See *op. cit.*).

Holistic grounding, however, is not limited to logic. Nor is it limited to the special grounding of logical constants delineated above. Nor is holism, as conceived here, restricted to the grounding project. Metaphysics in general deals with very basic issues, and a holistic methodology, modeled after foundational holism and

26 For details, see Sher (*op. cit.*)

holistic grounding, is especially suited for a substantive, highly explanatory discussion of such issues.

My answers to the questions "Where are you going, metaphysics?" and "How are you getting there?" are: "You are going where you have always gone, toward an objective and highly explanatory account of basic philosophical issues", and "To get there, you have to discard the traditional foundationalist, strictly-hierarchical methodology and adopt a new, flexible yet highly demanding methodology, a holistic methodology such as holistic grounding or its epistemic prototype, foundational holism."

Bibliography

Barnes, E. (forthcoming), "Symmetric Dependence", in *Reality and Its Structure*, edited by G. Priest and R. Bliss, Oxford.

Benacerraf, P. (1973), "Mathematical Truth", in *Journal of Philosophy*: 70, 661–679.

Bliss, R. and Trogdon, K. (2014), "Metaphysical Grounding", in *Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta, <https://plato.stanford.edu/entries/grounding/>.

Carnap, R. (1932), "The Elimination of Metaphysics Through Logical Analysis of Language", in *Logical Positivism* [1959], edited by A. J. Ayer, New York: Free Press, 60–81.

Chang, R. (2013), "Grounding Practical Normativity: Going Hybrid", in *Philosophical Studies*: 164, 163–187.

Coffa, A. (1979), "The Humble Origins of Russell's Paradox", in *Russell*: *the Journal of Bertrand Russell's Studies*: 33, 31–37.

Cummins, R. (1998), "Reflections on Reflective Equilibrium", in *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, edited by M. R. DePaul and W. M. Ramsey, Lanham: Rowman & Littlefield, 113–128.

Dasgupta, S. (2014), "On the Plurality of Grounds", in *Philosophers' Imprint*: 14, 1–27.

Denkel, A. (1996), *Object and Property*, Cambridge.

Dummett, M. (1973/81), *Frege*, New York: Harper & Row.

Fine, K. (2001), "The Question of Realism", in *Philosophers' Imprint*: 1, 1–30.

Fine, K. (2012a), "Guide to Ground", in *Metaphysical Grounding*, edited by F. Correia and B. Schnieder, Cambridge, 37–80.

Fine, K. (2012b), "What is Metaphysics?", in *Contemporary Aristotelian Metaphysics*, edited by T. E. Tahko, Cambridge, 8–25.

Friedman, M. (2001), *Dynamics of Reason*, Stanford: CSLI.

Grattan-Guinness, I. (1978), "How Bertrand Russell Discovered His Paradox", in *Historia Mathematica*: 5, 127–137.

Glymour, C. (1980), *Theory and Evidence*, Princeton.

Haack, S. (1993), *Evidence and Inquiry*, Oxford: Blackwell.

Harman, G. (1977), *The Nature of Morality*: *An Introduction to Ethics*, Oxford.

Hornsby, J. (1997), *Simple Mindedness*: *n Defense of Naive Naturalism in Philosophy of Mind*, Cambridge: Harvard.

Kant, I. (1781/87), *Critique of Pure Reason* [1929], translated by N. Kemp Smith, London: Macmillan.
Linnebo, Ø. (2008), "Structuralism and the Notion of Dependence", in *Philosophical Quarterly*: 58, 59–79.
Moore, G. H. (1988), "The Roots of Russell's Paradox", in *Russell*: *the Journal of Bertrand Russell's Studies*: 8, 46–56.
Quine, W. V. O. (1935), "Truth by Convention", in Quine (1976), 77–106.
Quine, W. V. O. (1954), "Carnap and Logical Truth", in Quine (1976), 107–132.
Quine, W. V. O. (1976), *The Ways of Paradox and Other Essays*, Cambridge: Harvard.
Rosen, G. (2010), "Metaphysical Dependence: Grounding and Reduction", in *Modality: Metaphysics, Logic, and Epistemology*, edited by R. Hale and A. Hoffman, Oxford, 109–136.
Schaffer, J. (2009), "On What Grounds What", in *Metametaphysics*, edited by D. Chalmers, D. Manley, and R. Wasserman, Oxford, 347–383.
Sellars, W. (1956), *Empiricism and the Philosophy of Mind*, Cambridge: Harvard.
Sher, G. (1991), *The Bounds of Logic*, Cambridge: MIT.
Sher, G. (1998), "On the Possibility of a Substantive Theory of Truth", in *Synthese*: 117, 133–172.
Sher, G. (1999), "Is Logic a Theory of the Obvious?", in *European Review of Philosophy*: 4, 207–238.
Sher, G. (2004), "In Search of a Substantive Theory of Truth", in *Journal of Philosophy*: 101, 5–36.
Sher, G. (2013), "The Foundational Problem of Logic", in *Bulletin of Symbolic Logic*: 19, 145–198.
Sher, G. (2016a), *Epistemic Friction*: *An Essay on Knowledge, Truth, and Logic*, Oxford.
Sher, G. (2016b), "Substantivism about Truth", in *Philosophy Compass*: 11, 818–828.
Sider, T. (2011), *Writing the Book of the World*, Oxford.
Simons, P. (1994), "Particulars in Particular Clothing: Three Trope Theories of Substance", in *Philosophy and Phenomenological Research*: 54, 553–575.
Sosa, E. (1980a), "The Raft and the Pyramid: Coherence versus Foundations in the Theory of Knowledge", in Sosa (1991), 165–191.
Sosa, E. (1980b), "The Foundations of Foundationalism", in Sosa (1991), 149–164.
Sosa, E. (1991), *Knowledge in Perspective*, Cambridge.
Sosa, E. (1997), "Reflective Knowledge in the Best Circles", in *Journal of Philosophy*: 94, 410–430.
Thompson, N. (2014), "Structuring Reality", *PhD Thesis*, University of Birmingham.

Benjamin Schnieder

On the Relevance of Grounds

Abstract: Three traditional philosophical issues that van Inwagen discusses in his metaphysical works are the Principle of Sufficient Reason, the question of why there is something rather than nothing, and the question of whether free will is compatible with determinism. The three topics are connected by a conceptual tie: the notion of a ground. In this essay, it is argued that van Inwagen's take on the three topics, ingenious as it otherwise is, suffers from an inadequate conception of the underlying notion of a ground.

1 Introduction

1.1 A Leibnizian web of topics

In the rationalist tradition spanning from Baruch de Spinoza to Arthur Schopenhauer, the so-called *Principle of Sufficient Reason* (henceforth: the *PSR*) was held in high esteem. G. W. Leibniz, who lent the *PSR* its title, even declared it to be one of two main principles of thought, alongside the principle of contradiction.[1]

One famous use that Leibniz made of the principle was to argue for the existence of God.[2] His argument was a cosmological one, starting with what he called the *First Question*: Why is there something rather than nothing? His reasoning, in a nutshell, was: As a matter of fact, the world is not empty but inhabited by things. The *PSR* tells us that there must be a reason for this fact. Such a reason

0 When I was an undergraduate, free will was one of the topics that I was drawn to. Whether or not incompatibilism is the folk view on free will (see Nahmias *et al.* (2006)), it was certainly *my* initial view on the issue. But then I read Schopenhauer, Moore, and Schlick, and I was hooked by their ideas of how to reconcile determinism with free will. By the time I came across Peter's Consequence Argument, I had become a steadfast compatibilist. So Peter's argument gave me a hard time. While its conclusion seemed clearly false to me, the argument was so neat and tidy, and I could not get my head around it. Only later, when I prepared for my finals, I had an idea of how to react to the argument. The idea became the material for one of my first publications (Schnieder (2004)). I guess one could say that I owe Peter for the start of my philosophical career. I therefore dedicate this paper to him in gratitude.

1 *Monadology*, §§31f.

2 *Principles*, §§7f.

https://doi.org/10.1515/9783110664812-005

cannot be part of the world itself because the reason is to account for *everything* in the world. So the reason must lie outside of the totality of contingent existents. It must therefore lie in a necessary existent which is its own reason and which created the world – in other words: God.

But although the *PSR* was held in high esteem in the rationalist tradition, it was also seen as beset with two related problems:

(i) If *everything* has a sufficient reason, then everything that *happens* has a sufficient reason. Determinism! So our feeling that our future is genuinely up to us is wrong – free will is an illusion.[3]

(ii) Even worse, any sort of contingency is an illusion. For, a sufficient reason necessitates that which it is a reason for. If everything is necessitated by a sufficient reason, then everything has to be as it is – so if the *PSR* is true, everything is necessary. There are no real alternatives to the world we live in. *Necessitarianism* rules.[4]

It certainly puts the *PSR* under quite some pressure if it has the said consequences. Some friends of the *PSR* have therefore tried to argue that the principle does not have such consequences after all. Others were willing to buy into determinism and even necessitarianism and downplay the costs of these claims. Leibniz, for instance, thought that the *PSR* indeed entails a form of determinism, but that free will, properly understood, is compatible with that sort of determinism.[5] And he held that although the *PSR* does entail that every truth is necessary *in some sense* of the word, there is an important alternative understanding of 'necessary' in which the *PSR* is compatible with there being contingent facts.

1.2 Van Inwagen on the Leibnizian web of topics

Enough about the rationalists. This is not an essay in the history of philosophy. It is an essay on topics – more precisely, on the *PSR* and three recurring topics from the rationalist debate about that principle: necessitarianism, Leibniz's First Question: Why is there anything at all?, and determinism and free will. That these perennial issues don't grow old is nicely illustrated by the fact that all of them figure prominently in the works of one of the most important metaphysicians of our times, Peter van Inwagen.

3 For this thought see, e.g., Crusius's *De Usu*, §§VIII, XXIII–XXIV.

4 For this thought see, e.g., Crusius's *De Usu*, §V.

5 On this and the following point, see Leibniz's *Discourse*, §13.

Although van Inwagen wrestles with a Leibnizian web of topics, his views on them are decidedly different from Leibniz's. Whereas Leibniz used the *PSR* in his cosmological argument for the existence of God, van Inwagen mounts a cosmological argument against the *PSR* – by arguing that the *PSR* indeed entails necessitarianism, which proves the principle to be wrong. Good for us, van Inwagen thinks. For, he acknowledges that the *PSR* may indeed entail determinism; and he takes his famous Consequence Argument to show that compatibilist accounts of free will are wrong. So the *PSR* would be a genuine threat to our freedom. As to the question of why there is anything at all, van Inwagen defends his own answer: There is only one empty world and countlessly many inhabited worlds. That makes it much more probable that the actual world is inhabited than that it is empty. But this answers Leibniz's question: There is something rather than nothing because the chances for there being nothing rather than something had been thin; infinitesimally thin indeed.

Van Inwagen's contributions to the debate about these topics are prime examples of good philosophical work. They are informed by traditional thoughts but at the same time fresh and original, they are rigorous, they are insightful. I truly admire them. Notwithstanding my admiration, I disagree with van Inwagen on all of them. Underlying my disagreement about the individual topics is a common core. Let me explain.

1.3 A core ingredient of the Leibnizian web of topics

The Leibnizian topics share a pivotal conceptual ingredient: the notion of a *reason*, or as I prefer to put it: a *ground*.[6] Grounds are what the *PSR* is about. Without any understanding of the concept *ground*, we can get no grip on the principle in the first place; without a *deepened* understanding of the notion, we cannot safely determine what consequences the principle may have. A ground furthermore is what a why-question asks for. Understanding Leibniz's First Question and evaluating proposed answers therefore requires an understanding of the concept *ground*; and deficiencies in our understanding of the notion can easily

6 In the pertinent literature, the terms 'reason' and 'ground' are used as translations for the same Latin ('ratio') and German ('Grund') terms. While the term 'reason' has become a standard in translations of the rationalists and in general in the discussion about the *PSR*, it is potentially misleading because it has a strong epistemic ring to it, while the notion expressed is not an epistemic one (which is why, e.g., Walford prefers 'ground' even in translations of rationalists talking about the *PSR*; see Kant, *Theoretical Philosophy* 485, note 5; on the non-epistemicity of grounds, see also below, §2).

misguide our judgements on the issue. Finally, the notion of a ground is also highly relevant to the Consequence Argument; but since this is less obvious, let me postpone this point until later, when I turn to the Consequence Argument.

The Leibnizian web of topics can only be adequately dealt with if we get a grip on the pertinent notion of a ground. Let me therefore turn to this notion now (see §2). While I cannot develop a full-blown theory of the notion here, I can present some cornerstones of such a theory that equip us for the discussion of the Leibnizian topics (which will follow in §3).

2 Grounds

Let me now try to shed some light on the notion of ground. No vulgar suspense: What I will not do, and indeed cannot do, is provide a *conceptual analysis* or *definition* of that notion. The notion may well be an indefinable primitive; in any case, I have nothing to offer on that front.

However, we understand many expressions whose meaning we cannot analyze. And we can deepen our understanding of expressions by putting them to use, delineating their correct applications, and formulating significant links to other expressions as well as principles which may be thought to be constitutive of their meaning.

So, what are grounds, in the sense of the word relevant to the Leibnizian issues?

First part of the answer: Grounds are facts and/or true propositions (henceforth, I will for simplicity's sake make no distinction between facts and truths).[7]

Second part of the answer: A fact is a ground in so far as it stands in a certain *relation* to another fact (following philosophical custom, we can call this relation *grounding*). The relation in question is one of productive priority: Some facts obtain *because of* others, their grounds. Grounds are *prior* to the facts they ground and *make* them obtain. The grounded facts, in turn, are *posterior* to their grounds and *depend* on them for their obtaining. To put the same thought in terms of truths: Some propositions are true *because* certain other propositions are. The former *owe* their truth to the latter, their grounds. The grounds are prior to the former and *account* for their truth. – The variety of locutions I used highlights

7 While I talk about grounding as a relation here, I am sympathetic to the idea that grounding may, *at bottom level*, better be conceived of in non-relational terms and best expressed in terms of a sentential connective (see Correia and Schnieder (2012b), pp. 10f.). But this issue will not play any role here.

that the relation of grounding is a *productive* relation. Grounds *bring about* what they ground. They are not merely accompanied by the latter; they are responsible for the latter. Note that the productivity of grounds need not be *causal* in nature. In fact, non-causal grounds are of particular interest in philosophy and will also be of central importance here.[8] To give just three examples here to illustrate the point: Physicalism (in one sense of the term) can be understood as the thesis that physical facts non-causally ground mental facts; the existence of a horse non-causally grounds the existence of the singleton containing that horse; and facts about the activities of an author non-causally ground facts about the fictional characters that occur in her stories.

Third part of the answer: Grounds are factors which bring about a fact (or: the truth of a proposition), *independently* of whether we *know* that they do this. Grounding is *not* an epistemic notion. Grounds must be distinguished from *evidence* and *reasons to believe*, and they must also be distinguished from *explanations*, at least if the term 'explanation' is taken in its ordinary meaning. For, whether we ordinarily call something an explanation depends on *epistemic* factors; an explanation should be enlightening with respect to some of our goals of inquiry. Grounding, however, is not an epistemic affair; whether a fact grounds another fact is independent of our goals of inquiry (just as it is independent of our goals of inquiry whether some event causes another). That notwithstanding, grounding may still be importantly related to explanations. Often, being told the grounds of a fact may essentially contribute to an explanation of the fact (just as being told the causes of an event may essentially contribute to an explanation of the event).

Fourth part of the answer: Grounds are finely individuated; finer than intensionally. They must be, because ground-theoretical vocabulary is hyperintensional: In the scope of ground-theoretical terms one cannot always substitute an expression with a co-intensional one *salva veritate*. To see this, take Aristotle's basic insight on truth:[9]

8 In the recent debate, talk about grounding is usually *restricted* to non-causal cases. In the rationalist tradition, however, philosophers often countenanced a notion of ground that covers causal and non-causal cases alike (Bolzano – in his 1837: book II, §§ 168, 192 – even went one step further and tried to analyse the concept of causation in terms of a non-causal notion of ground; for a discussion of his proposal, see Schnieder (2014)). I am in general open to that idea, though a lot depends on how exactly it is spelled out. But a discussion of the issue would lead us too far astray here. (For some important thoughts on the issue, see Schnieder (2016).)

9 See Aristotle *Metaphysics*, book Θ 10: 1051^b6–9. The Aristotelian Insight plays an important role in the contemporary debate about truth and truth-making; see e.g. Künne (2003), ch. 3.5, and Rodriguez-Pereyra (2005).

INSIGHT That snow is white is true *because* snow is white.

The connective 'because' is one of the linguistic devices used to introduce the relation of grounding: That snow is white is a ground of its being true that snow is white. Now take a look at the two clauses connected by 'because' in sentence INSIGHT, i.e. 'snow is white' and 'that snow is white is true'. They are co-intensional: They are true with respect to exactly the same possible worlds. However, substituting (in INSIGHT) one for the other generates falsities:

INSIGHT? Snow is white *because* snow is white.
INSIGHT?? Snow is white *because* it is true that snow is white.

The first of these claims is false because the fact that snow is white does not bring about itself (I think no fact does; but I am *certain* this one doesn't). The second of the claims is false because it gets the order of things wrong. In a slogan: truth depends on being, not vice versa.

The Aristotelian Insight therefore illustrates the hyperintensionality of ground-theoretic vocabulary such as 'because' or 'is a ground of'. That means that the relata of grounding, i.e. facts or truths, must be finer individuated than intensions. They cannot, for instance, be sets of possible worlds. If they were, then whatever is grounded by the fact that snow is white would also be grounded by the fact that it is true that snow is white; for it would be one and the same fact.

For another example, consider the disjunction that grass is green or some equiangular triangles are not equilateral. The disjunction is true. Why? Because grass is green. Now notice that the disjunction and its first disjunct are true in exactly the same worlds. Still, the disjunct grounds the disjunction and not vice versa. Another class of examples which illustrates the need for finely individuated grounds played a crucial role in the rationalists' debate: hierarchies of mathematical truths.[10] Mathematical truths are not all on the same level; some are more fundamental and ground others. Basic axioms of arithmetics ground complex theorems, for instance. However, since mathematical truths are necessary truths, they are all necessarily co-obtaining. So if we individuated mathematical truths intensionally, there would be only a single one. But then we could not allow that some mathematical truths ground others. To mention one last class of examples, take metaphysical truths. Given that such truths are necessary truths, there would be only one of them if we individuated them intensionally (and that single metaphysical truth would also be the single mathematical truth). But some metaphysical

10 See e.g. Bolzano, *Beyträge*; Schopenhauer, *Wurzel*, ch. 6.

truths ground others. Assume for the moment that, by necessity, murdering innocent human beings is wrong; and assume that, by necessity, there is a divine being which abhors the murdering of innocent people. As Socrates reminded us in the *Euthyphro* dialogue, we may still reasonably wonder whether murder is wrong because the divine being abhors it, or whether the divine being abhors murder because it is wrong. But we are dealing, per assumption, with necessities here; so the question about which of the two facts grounds which requires that we do not individuate facts intensionally.

Fifth and final part of the answer: Grounds like to cooperate. Often, it is not a *single* fact that grounds another fact; instead, two or more facts *jointly* do the job. What grounds the fact that Obama has two daughters? It is jointly grounded by three facts: the fact that Malia is Obama's daughter, the fact that Natasha is Obama's daughter, and the fact that Malia and Natasha are distinct. So, grounding is not a dyadic relation between facts; instead, it has a variable adicity with respect to the position of the ground.

Summing up the above, grounds are facts/truths with four important features:

Productivity	Grounds bring about other facts/truths.
Non-epistemicity	Grounding is not relative to epistemic factors such as our goals of inquiry.
Finegrainedness	Since ground-theoretical vocabulary is hyperintensional, grounds must not be individuated intensionally.
Cooperativity	Two or more facts can work together and *jointly* ground another fact.

There is much more to say about grounds and the relation of grounding.[11] But the features of Productivity, Non-epistemicity, Finegrainedness, and Cooperativity suffice to get our discussion going.

3 Back to the Leibnizian topics: The PSR, the first question, and free will

I will now examine van Inwagen's views on the Leibnizian issues introduced earlier: the *PSR*, the question 'Why is there something rather nothing?', and the (in)compatibility of free will and determinism. I will argue that van Inwagen's central arguments, ingenious as they are, should be resisted. The main reason is that he underestimates the relevance of grounds; that is, while the notion of a ground plays a pivotal role for all the issues, van Inwagen downplays its im-

11 For more detailed expositions that I endorse see Rosen (2010) and Fine (2012).

portance. He avoids a close and careful examination of the notion; as a result, the arguments misconstrue aspects of the nature of grounds. In particular, the arguments often rely on a modal framework. The great success of modal logics in the second half of the twentieth century made this framework a natural choice for van Inwagen to work with. Modality was a well understood notion, and for many purposes, a modal framework indeed works well. And yet, it is too blunt a tool to adequately deal with issues of ground.

The following discussion revives arguments from previous publications of mine (some of which I co-authored with colleagues) and puts them into a larger context. Thus the content of §3.1. draws on Schnieder and Steinberg (2016), the content of §3.2. draws on Kappes and Schnieder (2016), and the content of §3.3. draws on Schnieder (2004). Each of those three papers dealt with only one of the issues I will address in what follows, which gave me more space for the individual topics. I will therefore concentrate on the bigger picture here and brush over some details (for which I would refer the reader to my previous publications).

3.1 The PSR and necessitarianism

The PSR says:

PSR Every truth has a sufficient reason/ground.

That a ground is sufficient means (in the present context) that the ground *necessitates* what it grounds. In schematic form:[12]

Sufficiency
<p> is a sufficient ground of <q> $\stackrel{df}{\equiv}$ (i) <p> is a ground of <q> & (ii) □(if p, then q).[13]

Van Inwagen thinks the failure of the *PSR* can be demonstrated by the use of a cosmological argument which shows that even if most propositions could have a sufficient reason, there still remains one which can't, namely the conjunction of all contingent propositions (call it *The Big Contingency*; for short: *BC*). Van Inwagen argues that this proposition constitutes a counterexample to the *PSR*, at

12 In the rationalists' debate about the PSR, the term 'sufficient' is used in a number of different meanings, including a modal one such as the one above (see Crusius's *De Usu*, §III). In the contemporary debate, the modal meaning is the predominant one.

13 I use pointed brackets to refer to the fact/truth expressed by the bracketed expression.

least if there are any contingent propositions. Since there obviously are contingent propositions, the *PSR* fails.

For his argument, van Inwagen presupposes a restricted version of the irreflexivity of grounding:

> Irreflexivity* If x is a contingent truth, then x is not a ground of itself.[14]

The argument then runs as follows:

> Let P be the conjunction of all contingently true propositions into a single proposition. [...] It is evident that P itself is a contingent proposition, for a necessary proposition may not have even a single contingent conjunct. Now, according to *PSR*, there exists a state of affairs S that is a sufficient reason for P. S must be contingent or necessary. But it cannot be either. It cannot be necessary, for, if it were necessary then P (which, by [Sufficiency], is entailed by S) would be necessary. It cannot be contingent, for if it were contingent, it would be a conjunct of P; and if it were a conjunct of P it would be entailed by P; and if it were entailed by P, it would both entail and be entailed by P; and if it both entailed and were entailed by P, it would *be* P (by the criterion of identity for states of affairs given [earlier]); and if S *were* P, then a contingent state of affairs would be its own sufficient reason, contrary to [Irreflexivity*]. Since S cannot be either necessary or contingent, it cannot exist, and *PSR* is false.
>
> This result follows provided we assume that there is such a thing as P – that is, the conjunction of all contingently true propositions. And there is such a thing as this if there are *any* contingently true propositions. Hence if *PSR* is true, there are no truths but necessary truths: there is no distinction to be made between truth and necessity. We must therefore reject *PSR* [...]. (Van Inwagen (1983), pp. 203f.)

The argument is clear as it stands; further explications seem unnecessary.

I am not a friend of the *PSR*. I do not believe in the principle; I even believe it is false. It would be nice to see a proof that vindicates my belief. But in my view, van Inwagen's argument is no such thing. The reason is that it rests on an inadequate conception of the *relata* of grounding. Earlier, I pointed out that ground-theoretical vocabulary is hyperintensional. Therefore, grounds and what they ground, i.e. the relata of the grounding relation, must be finer individuated than intensions; they must allow that *two distinct* truths are necessarily equival-

14 That grounding is an irreflexive relation is a commonly shared assumption in the current debate (Raven (2013) calls it part of the orthodoxy in the debate). Van Inwagen (1983), p. 203, employs only a restricted version of irreflexivity to make room for the view that necessary truths are grounded in themselves, without himself committing to it.

ent, i.e. that they are true in the same possible worlds.[15] But the central move of van Inwagen's argument presupposes an intensional individuation of grounds: It says that if the *Big Contingency BC* were grounded in one of its conjuncts, there would be a mutual entailment between the conjunct and *BC*, *and therefore* the conjunct would be identical to *BC* (violating the irreflexivity of grounding). This is a non-sequitur; we see this once we realize that grounds are not individuated intensionally. As it stands, the argument therefore fails to establish its conclusion.

It makes sense to wonder, though, whether the argument's marriage to the intensional individuation of grounds is essential for it, or whether the argument can be salvaged in a modified form. There is indeed a version of the argument that does not explicitly presuppose the intensional individuation of grounds. Jonathan Bennett puts it as follows:[16]

> Let *P* be the great proposition stating the whole contingent truth about the actual world, down to its finest detail, in respect of all times. Then the question 'Why is it the case that *P*?' cannot be answered in a satisfying way. Any purported answer must have the form '*P* is the case because *Q* is the case'; but if *Q* is only contingently the case then it is a conjunct in *P*, and the offered explanation doesn't explain; and if *Q* is necessarily the case then the explanation, if it is cogent, implies that *P* is necessary also. (Bennett (1984), §25, p. 115)

In his version of the argument, Bennett makes no mention of any particular take on the individuation of grounds. What he presupposes is that we cannot explain a conjunction by one of its conjuncts. This assumption is also made by Ross and Rowe, who had proposed versions of the argument before Bennett and van Inwagen proposed theirs.[17] Rowe justifies the assumption as follows: If a proposition

> '*q* explains [the *Big Contingency*, *BC*], then *q* cannot be contingent; otherwise it would be a part of [*BC*] and [*BC*] (a contingent state of affairs) would be self-accounting, which is impossible.' (Rowe (1975), p. 107)

But why would *BC* be *self*-accounting if it is accounted for by one of its *conjuncts*? Van Inwagen's version of the argument can be seen as a failed attempt to make this claim plausible. But what else might stand behind the claim? The idea may simply be that conjunctions ground their conjuncts; so if *BC* were grounded in one

15 Van Inwagen presents a variation on the argument in his *Metaphysics* (Van Inwagen (2009), pp. 150ff.). The variation also presupposes a framework in which grounds are intensionally individuated and thus fails for the same reason.

16 Van Inwagen reported to me that he and Bennett actually worked together on the argument.

17 See Ross (1969), Rowe (1975). Ross's version is the first I could find in the literature, but there may well be predecessors.

of its conjuncts, we'd have a circle of grounds; if we can moreover chain grounding claims, *BC* would then indeed be self-accounting.[18]

But the idea that conjunctions ground their conjuncts should be resisted; once one starts reflecting on the idea, one has to realize its odd consequences. Imagine a toy world, with two fundamental facts, *P* and *Q*. It would be silly to insist that there is a third fundamental fact, namely the conjunction of *P* and *Q*. Instead, what one should obviously say is that the conjunction is a derivative fact, a grounded one: the two fundamental facts *P* and *Q* make their conjunction obtain. Moreover, the idea that conjunctions ground their conjuncts immediately leads to the odd consequence that *all* truths are grounded: For, every truth is a conjunct in some conjunction. That would be happy news for the friend of the *PSR*.

Conjunctions do not ground their conjuncts; in fact, this claim turns things upside down. This was suggested by a reflection on the toy world example. But it is also supported by a very general thought: Propositions governed by classical logical connectives have their truth-values *because of* the truth-values of their components.[19] A disjunction with one true disjunct, for instance, is true because the disjunct is. The fact that Socrates is wise grounds the disjunctive fact that Socrates is wise or lazy. And a true conjunction is true because of its two conjuncts, which *jointly* make the conjunction true. Hence, the *Big Contingency* is grounded; its grounds are its conjuncts.[20]

An objection that is sometimes voiced against the idea that conjunctions are grounded in their conjuncts is phrased in terms of explanations. Take a particular conjunctive truth, e.g. that a certain disc is flexible and heavy. If we ask why the disc is flexible and heavy, we will clearly not be content by being told that it is so partly because it is flexible, and partly because it is heavy. So it may seem that we intuitively reject the idea that conjuncts ground their conjunction.

However, this objection confuses *grounding* and *explanation*.[21] Talk of explanation has a strong epistemic ring to it, which easily distorts the issue. When we

18 The transitivity of grounding is commonly assumed in the current debate; Raven (2013) calls it part of the orthodoxy and discusses problems raised by Schaffer (2012); for another reply to Schaffer see Krämer and Roski (2017).

19 This is a core idea of all current approaches to the propositional logic of grounding; see e.g. Correia (2010), Schnieder (2011), Fine (2012).

20 Vallicella (1997) brought forth the same sort of criticism against van Inwagen's argument, as did Rowe (Rowe (1975), pp. 107f., fn. 32) against Ross's version of the argument. A certain disadvantage of their criticisms is that it is phrased in terms of the *explanation* of a conjunction; I will show in a minute why that may distort the issue.

21 On the following compare also Schnieder (2016).

ask for explanations, we have certain explanatory interests; the interests we typically have when we ask for an explanation of a conjunctive truth will presumably be to know why its conjuncts are true. Hence, being told that the conjunction is grounded in its conjuncts will not meet them. But this is not an argument against conjunctions' being grounded in their conjuncts. We reject the answer not because it is false, but because it is not the *sort* of answer we wanted to hear. Thus, the objection only highlights a fact about our contingent explanatory interests, not a fact about how conjunctions are grounded.

Conjunctions do not ground their conjuncts. To the contrary, they are grounded in them. Let me mention two consequences for the *PSR*:

Firstly, our initial formulation of the *PSR* said that every fact *has a sufficient ground*. How should we understand the italicized phrase here? On a reasonable reading, to have a ground here means to *be grounded*; since grounds are cooperative, that does not necessarily mean to be grounded in a *single* fact. Several facts may jointly do the effort. One problem of the cosmological arguments against the *PSR* is that they overlook the cooperativity of grounds and only take into account grounds which are single facts.

Secondly, once it is acknowledged that conjunctions are grounded in their conjuncts, it turns out that we have not only rebutted Rowe's particular version of the argument. For we now see that the very idea of the argument is on the wrong track: What all the versions of the argument have in common is that a huge conjunction is chosen as a counterexample to the *PSR*.[22] But any true conjunction *is* grounded, namely in its conjuncts. Hence we know that conjunctions are the wrong *sort* of proposition to constitute counterexamples to the *PSR*. This leads to a more general upshot of the discussion. It concerns where we should look for possible counterexamples to the *PSR*. The cosmological argument against the *PSR* looks at a particularly 'big' thing (a huge conjunction) that is constituted by 'smaller' ones (its conjuncts). But, generally speaking, the properties of big things are grounded in those of smaller things; so the argument looks in the wrong place. If there are counterexamples to the *PSR*, they will be found among the particularly

22 Incidentally, there are good reasons to doubt that there can be such a thing as the conjunction of all contingent truths, unless one works with intensionally individuated propositions. Since the conjunction of all contingent truths is itself contingent, it would have to have itself as a conjunct. But if one thinks that propositions are *structured* and have parts, one should be suspicious about the existence of a conjunction which is one of its own conjuncts, since this conjunction would be its own proper part. But even if one conceives of propositions not as mereologically structured, we should deny the existence of a conjunction that has itself as a conjunct once we accept that the conjuncts of a proposition jointly ground the proposition; the said conjunction would then be grounded in itself and thereby violate the irreflexivity of ground.

small things. Logically simple facts about tiny particles, for instance; or basic axioms about numbers. Grounding runs from the simpler to the complex. Hence, to find something ungrounded, one has to turn away from the complex things and turn to the simple ones.

Summing up, I argued that van Inwagen's argument against the *PSR* fails (as do similar cosmological arguments). The argument could seem compelling as long as the notion of a ground remained in the dark; in particular, as long as some important features of grounds were neglected, in particular their Finegrainedness and their Cooperativity. Moreover, the fact that conjunctions are grounded in their conjuncts could be obscured by ignoring the non-epistemicity of grounds and confusing grounding with explanation.

3.2 The first question: Why is there anything at all?

Turning to Leibniz's First Question, let me recapitulate van Inwagen's take on the issue: There is at most one empty world, whereas there are infinitely many populated worlds. But all worlds have the same probability of being actualized. Hence it is immensely improbable that there is nothing. And this answers why there is something: There is something because it was extremely probable that there be something.[23]

While van Inwagen thinks this is a good answer to the First Question, he thinks it is not the *best* sort of answer one could hope for; but he also thinks that unfortunately the best sort of answer is unavailable. What would it look like?

> One sort of answer, the best if we could get it, would consist in a demonstration that it was impossible for there to be nothing. Or so I would suppose: if showing that it is impossible for a certain state of affairs to obtain doesn't count as answering the question why that state of affairs does not obtain, I don't know what would count. (Van Inwagen (1996), p. 95)

The question 'Why is there something?' is a request to be informed about the *grounds* of the truth that there is something. Van Inwagen thinks that if we show that some fact necessarily obtains, we thereby reveal why the fact obtains. In other words, he seems to subscribe to the following principle:

23 See Van Inwagen (1996). For a critical discussion of how van Inwagen reaches his conclusion, I recommend Rodriguez-Pereyra's 'Why is there something rather nothing? A probabilistic answer examined' (he presented this paper at the conference Van Inwagen (1996) in Warsaw held in honour of van Inwagen).

NECESSITY AS A GROUND
If $\Box p$, then (p because $\Box p$).
In other words: If $\Box p$, then the fact that $\Box p$ grounds the fact that p.

But I find this principle highly implausible. If the question 'Why is there something?' is a request to be informed about the *grounds* of the truth that there is something, the answer 'Because it is so by necessity.' is no good. Recall that grounds are *productive*; they are prior to what they ground and bring it about. But at least to my eyes, it seems off to suppose that the modal status of a fact (that it is necessary) brings about the fact itself. It is the wrong sort of thing to do that.

Here I anticipate the objection that I just voiced my personal intuition to the effect that NECESSITY AS A GROUND is false. So let my try to lend some support to it.

Here is a more general thought that may back the intuition: The modal status of a fact is one of its properties; but how should a property of a thing bring about the existence of the thing? It seems that the thing must already exist in order to have a property. Let me admit, however, that this thought depends on two assumptions, namely that grounds are *facts* (rather than true propositions), and that for facts, to exist is to obtain. Since I myself would not commit to these assumptions, I would not put much money on this defence of my intuition.

However, I think the implausibility of NECESSITY AS A GROUND becomes more obvious once one considers the fact that necessity reiterates. If it is necessary that p, then it is also necessary that it is necessary that p ($\Box p \rightarrow \Box\Box p$). But given some necessary fact, it is a non-trivial question as to *why* the fact is necessary, i.e. what *makes* the fact necessary. Answers will presumably vary, depending on the kind of case in question. Some facts may be necessary because of some *conceptual* relations (an example *might* be the fact that bachelors are unmarried). Others because of *non-conceptual*, *metaphysical* factors (an example *might* be the fact that Socrates is human if he exists at all). Still others may perhaps inherit their necessity from other facts (an example might be the fact that $\Box\Box 2 = 2$). But equipped with NECESSITY AS A GROUND, we can end the discussion. Given *any* necessary fact, we already know what makes it necessary: that it is necessary that it is necessary. It is all too easy account to for the necessity of a fact – oh happy metaphysician!

NECESSITY AS A GROUND should not be accepted. I will not go into detail here, but I daresay that if NECESSITY AS A GROUND fails, we have good reasons to suspect that the analogous principle PROBABILITY AS A GROUND (if probably p, then the fact that probably p grounds the fact that p) will do so all the more. Hence, while

I find van Inwagen's answer to the First Question ingenious, I also find it false. That a fact is very probable is not a ground of the fact.[24]

So far for my reason why I am not content with van Inwagen's answer to the First Question. I am not completely sure, however, whether we are on the same page with respect to what is at stake. Leibniz's question is a why-question. Such a question can sometimes be understood epistemically. It can be a request for *justification* in favour of its topic ('Clearly, she likes him.' – 'Why?' – 'Cause she typed his script for him.'). This reading of the question makes hardly sense in the present case; we don't need justification for believing that there is something. We just have to open our eyes. However, a why-question may perhaps also be a request for another kind of epistemic favour. It might be a request to provide information relieving us of some intellectual *surprise* we feel about a given fact, and/or information giving us the intellectual feeling of better *understanding* the issue in question. We may be surprised that something happened, because we did not expect it; or because we *would not* expect it, once we start reflecting on it. And this is indeed how some philosophers look upon many facts of the common and garden variety, including the fact that there is something. Once they start reflecting on such facts, they feel surprised and puzzled by them (whether this is the philosophers' gift or curse can be left open here). Here, van Inwagen's answer to the First Question could indeed be helpful; as could a demonstration that it is necessary for there to be something. For, being shown that a given event was very probable, or that a given fact obtains by necessity, can do an important epistemic job and remove the intellectual surprise someone feels about it. Still, this does not mean we now know the ground of why the event took place, or why the fact obtains. And although the question 'Why is there something rather than nothing?' can be read as a request for epistemic help, this reading is not the most crucial one here. First and foremost, the question asks about the *ground* of why there are things. At least this is what Leibniz meant by it when he introduced the question into the debate (in his *Principles*, §7), and some philosophers in his wake clearly meant the same by it.[25] Perhaps, though, not all of them do. Should this not be what van Inwagen means by it, then I do have no complaints about the answer he proffers; my only complaint then would be that he might have warned us that he changed the topic from the one Leibniz once introduced.

24 Here I am in agreement with Rodriguez-Pereyra (manuscript) (though we disagree with respect to the question whether a necessary fact can be grounded in its own necessity; unlike me, Rodriguez-Pereyra thinks this is sometimes the case).

25 See e.g. Rescher (2016).

So far, then, I argued that whereas van Inwagen's answer may be a helpful reply to the First Question understood in an epistemic fashion, it does not reveal the ground of why there is something. Thus it does not answer the question as originally meant by Leibniz: Why is there something rather than nothing, i.e., what is the ground for there being something rather than nothing?

Do I have an answer on offer? I think I do. Unfortunately, I am pretty certain that van Inwagen will not like it (for sure, Leibniz would not like it). But let me at least wrap the answer in some flattery: There is something rather than nothing because there is van Inwagen.

Admittedly, the flattering touch of the answer only goes so far. By answering that way, I do not mean to attribute any divine attributes to van Inwagen. Nor does he play any special role for the sort of answer; any old entity there is would have done just as fine. What grounds an existential fact, i.e. a fact expressed in terms of the existential quantifier, are its obtaining instances.[26] Imagine a toy world in which God creates Quine and turns him into a philosopher. It would be silly to think that God thereafter has something more to do to bring about the fact that there is a philosopher. This existential fact is a derivative one; it is grounded in the fact that Quine is a philosopher. A more general reason for accepting the view that existential quantifications are grounded in their true instances stems from observing that existential facts have a kind of disjunctive nature. They can be thought of as equivalent to the huge disjunction of their instances (even though they *are* not huge disjunctions, for reasons having to do with infinity and with topicality; but I will not go into this here). Just as the disjunction of their instances is grounded in its true disjuncts, the existential quantification is grounded in its true instances.

The question 'Why is there something rather than nothing?' has sometimes been called the deepest question there is. If my reasoning above is correct, that seems a tad exaggerated. There are oh so many simple answers to the question. So many that it seems hard, indeed, to think of any cheaper question.

But wait a second. As I said before, it seems clear that Leibniz would not have been content with this answer; nor would many others who took the question seriously. One reason is precisely that they thought the answer to the question *must* be a very profound one, while the answer I offer is a banal outflow of the logic of grounding. But why *did* they expect something deep in the first place?

My suspicion is they expected something deep because they had a slightly different question in mind. One often encounters the following formulations of the question: Why is there anything *at all*? Or: Why is there anything *in the first place*?

26 See, e.g., Rosen (2010), p. 117; Correia and Schnieder (2012b), p. 18; Fine (2012), p. 59.

At first glance, the italicized phrases may seem to be mere rhetorical ornaments. But I don't think they are; to the contrary, I think they make a huge difference to what's at stake. 'Because there is van Inwagen' is no appropriate answer to these questions. The reason is not, however, that the answer is *false* and that existential facts aren't grounded in their true instances after all. The reason is rather that this is not the *sort* of answer that the inquirer is after, and that the phrase 'at all' is a conventional indicator that a particular type of answer is requested. It is not easy to say exactly what type that would be. But roughly speaking, it seems that one is looking for an explanation of a phenomenon that does not itself invoke any instances of that phenomenon.

Why are there so many rabbits in these woods? Because there have been rabbits here for many years, and rabbits reproduce at a very high rate. But why were there rabbits here *in the first place*? The answer that was good before is no good here; an answer is required which does not invoke any rabbits being in these woods. An explanation citing some fact that is prior to the existence of any of the rabbits that have ever been in these woods. Why is there something rather than nothing (i.e. what is the ground of the existential fact that there is something)? Among other things, that there is van Inwagen. But why were there things *in the first place*? The answer that was good before is no good here; an answer is required which does not invoke there being any things at all.[27]

I for my part have no idea how such an answer that is free of any existential presuppositions might look like. And I do not think there is any reason why we should think there is such an answer. I don't believe in the *PSR*; hence I am open to the possibility that some why-questions don't have any true answers. But even if the *PSR* were true, this would only require there to be a ground of the fact that there is something. And we know that there are many grounds of that. The *PSR* would not require there to be a ground for that fact which does not invoke there being any entities in the first place.

Let me take stock. First upshot of the discussion: When people ask 'Why is there something rather than nothing?' they may request different types of answer. The question can be understood epistemically; then van Inwagen's answer seems a good step forward (and showing that there *has to be* something would take us

27 A historical aside: It seems that when Leibniz (*Principles*, §7) asked 'Why is there something rather than nothing?', he implicitly restricted the quantifier to contingent entities. What was to be explained then, was not the existence of entities as such, but rather the existence of some particular *sort* of entity. Leibniz's answer then indeed met the requirement of not invoking the phenomenon to be explained, since it proceeded in terms of a necessary substance and did not invoke any contingent entities.

even further). But it may be understood as a question about what *grounds* the existential fact that there are things. Then van Inwagen's answer is on the wrong track; facts are not grounded in their own likelihood (nor in their own modal status). Rather, what grounds an existential fact are its true instances. So, there are numerous grounds for the fact in question; as many as there are things. Some philosophers who ask the question aiming at a ground, however, implicitly place some additional constraints on the answer they want to hear; constraints that the answer I offered may well fail to satisfy. But truths stay true even if we are not interested in hearing them, so I do not think that the desire to see a particular sort of answer can speak against the claim that the existential fact is grounded in its true instances. Moreover, those who place constraints on the answer they want to hear should better make clear what exactly they are after; otherwise, we will all too easily talk past each other. And if the constraint is for the answer to be free of any existential presuppositions, the search may well be in vain. At least I see no reason why there should be such a thing.

3.3 Free will and the consequence argument

Is free will compatible with determinism?[28] Van Inwagen's Consequence Argument presents a very strong case for the negative answer.[29]

The argument makes use of the following principles:

> Determinism
> If determinism is true, then for all times t, t^* (with $t < t^*$):
> Any true state description of the world at t^* follows from a true state-description at the world at t, together with the laws of nature.

28 As noted earlier, the *PSR* seems to entail determinism, which is why freedom of the will was a recurring topic in the traditional debate about the *PSR*. This is also why van Inwagen discusses the *PSR* in his work on free will. (Let me note in passing, though, that Van Inwagen (1983), p. 202, is somewhat cautious about the claim that the *PSR* entails determinism. I think that given what else he believes, he should not be. For, he takes the *PSR* to entail necessitarianism. But if necessitarianism is true, every true state-description of the world at an arbitrary time is necessarily true, as are the laws of nature. Since all necessary propositions entail one another, every true state-description of the world at some time will then be entailed by any earlier true state-description of the world; thus we have determinism.)

29 See Van Inwagen (1975, 1983) (in his *Essay on Free Will*, van Inwagen presents three versions of the argument; I concentrate on the first of them here). My presentation slightly differs from van Inwagen's, but my criticism does not depend on that.

Free Will
For all agents A: If A has free will, then A had at some times the ability to act differently from how A actually acted.

Transference
For all agents A and all propositions p, q:
If A is able to render p false and p follows from q, then A is able to render q false.

The Consequence Argument is intended to show that if determinism is true, nobody ever had the ability to act otherwise. Hence, nobody ever had free will. For the argument, we consider an ordinary arbitrary agent A, an arbitrary time t, and an arbitrary action α the agent performed at t. Now we reason as follows:

P.1 If A, at t, was able to act otherwise, A was able to render it false that α occurs at t.
P.2 If determinism is true, then the conjunction of a true state-description of the world 1.000 years before A's birth and the laws of nature, entails that α occurs at t.
P.3 A was never able to render the conjunction of a true state-description of the world 1.000 years before A's birth and the laws of nature false.
C If determinism is true, then A, at t, was not able to act otherwise.

Once the principles Determinism and Transference are accepted, there is no way around the conclusion of the argument. P.1 is simply a description of our setting in terms of the phrase 'render something true'. P.2 directly follows from the setting together with principle Determinism. P.3 seems an obvious truth, given that we are dealing with an ordinary agent (and not with almighty God). The conclusion then follows by the aid of principle Transference.

Since we make no special assumptions about our trio of agent, time, and action, the result of the argument generalizes to all such trios. So, if determinism is true, nobody ever has the ability to act otherwise. So, by principle Free Will, if determinism is true, nobody enjoys free will.

This is a beautiful piece of philosophical reasoning. Yet, I am not convinced.

A crucial locution employed in the argument is that of an agent's being able to render a proposition true, which links talk about abilities to talk about the truth/falsity of propositions, and thereby to the characterisation of determinism. Even though the locution may sound a bit stilted, it is clearly well understandable. But can we say in detail how we understand it?

Any analysis of the locution should meet certain criteria of adequacy:

- In order to understand when someone is *able* to render something false, we must understand what it means to *render* something false (the description of the ability involves a description of an action).
- Any good analysis of rendering something *false* should allow for a strictly parallel analysis of rendering something *true*.

- Any good analysis should account for the following abilities and inabilities:
  - Any agent A who is able to perform the action of ϕ-ing at t, is also able to render it true that A ϕs at t (this constraint then gives us premise P.1 of the argument).
  - No human agent is able to render the laws of nature false.
  - No human agent A is able to render a true (false) state-description of the world at a time before A's birth true (false).
  - No human agent is able to render mathematical truths true, or mathematical falsities false. Mathematical propositions are beyond the reach of any human agents (in fact, I think, of any agents at all).[30]

I propose the following analyses:

> Agent A renders proposition P true (false) $\stackrel{df}{\equiv}$ A does something such that P is true (false) because A does it.
> Agent A is able to render proposition P true (false) $\stackrel{df}{\equiv}$ A is able to do something such that, if A did it, P would be true (false) because A does it.

The analysis meets all the desiderata.[31] It deals with truth and falsity in a strictly parallel fashion. It does not bestow agents with the incredible abilities of rendering laws of nature false, state-descriptions of remote times false, or mathematical propositions true or false. But it lets agents keep the abilities they should have: If Jean can sing, she can render it true that she's singing. For, she can sing; and if she sang, then it would be true that she sings because she sings (by the Aristotelian Insight on truth; see §2 above).

30 In my previous publications, I stated the requirement in more general terms and wrote that no human agent can render necessary truths true, or necessary falsities false. I overlooked a peculiar sort of case (thanks to Florian Fuchs for pointing this out): Assume Jean is singing; she thereby renders it true that she is singing. But the fact that she is singing grounds the fact that she is singing or isn't singing (disjunctions are grounded in their true disjuncts). If grounding is transitive (which I believe), it follows that Jean can render some necessary truths true (any instance of the excluded middle one of whose disjuncts she can render true). So some necessary truths are grounded in contingent truths, and moreover rendered true by human agents (incidentally, *any* classical tautology is grounded in some contingent fact, if standard rules for the logic of grounding are accepted; see Schnieder (2011), §4, for a proof and some discussion). This is I now believe that my earlier requirement was too strong. Still, *most* necessary truths cannot be rendered true by human agents. The above requirement picks out one class of them. But of course, there are others as well.

31 Unlike analyses previously offered in the literature, e.g. by Lewis (1981), p. 297, or Van Inwagen (1983), p. 68.

A crucial element in this analysis is the connective 'because' – hence the notion of a ground rears its ugly head again. Buried in the notion of rendering something false we find an element that introduces a hyperintensional aspect into our vocabulary.

This makes the locution less perspicuous than it may initially seem, and it makes the evaluation of principles involving it harder than it may initially seem. In fact, as I now will argue, one of the principles employed in the argument turns out to be false, now that we've become clear about what it means to render propositions true/false.

The principle in question is TRANSFERENCE. This principle certainly looks plausible at first; indeed, van Inwagen writes about it:

> This principle seems to be analytic. For if Q entails R, then the denial of R entails the denial of Q. Thus, any condition sufficient for the falsity of R is also sufficient for the falsity of Q. Therefore, if there is some condition that S can produce that is sufficient for the falsity of R, there is some condition (that same condition) that S can produce that is sufficient for the falsity of Q. (Van Inwagen (1975), p. 192)[32]

While I agree with van Inwagen that the principle seems plausible at first glance, perhaps even analytic, I nevertheless think it is false. Here is a counterexample: Any proposition whatsoever follows from a contradiction (given that we are working with classical logic). So if the principle were true, then any agent who can render something false, can also render false that 2 is prime and not prime. But nobody ever was in a position to render that false. The proposition is false, for sure; but its falsity was never up to us.

To consider another counterexample to TRANSFERENCE, assume that Jean, at a particular time t, is singing. She thereby renders it false that she is silent at t. But now consider the following two propositions:

> HISTORY
> 1.000 years before t, the total mass of matter in the universe was lower than 10 tons.
>
> LAW
> Once a system reaches a state with a total mass of matter lower than 10 tons, the matter will dissolve into energy at a constant speed of 2 tons per decade, until the system eventually reaches a terminal state free of any matter.

Together, the two propositions entail that Jean is not singing at t (given that only substances made of matter can sing; if you do not believe this, just modify the

32 Compare Van Inwagen (1983), p. 72.

example in an appropriate way). So the TRANSFERENCE principle entails that Jean could render the conjunction of HISTORY and LAW false. But she could not; what was the case 1.000 years ago is wholly out of her control, as are the laws of physics. Hence, the TRANSFERENCE principle fails.

Since the principle is false, something must have gone wrong with van Inwagen's defence of it. Here is my diagnosis: If an agent produces a condition that is sufficient for the falsity of p, and p follows from q, then she produces a condition that is *sufficient* for the falsity of q. But it does not follow that q is false *because* of this condition. And therefore it does not follow that the agent thereby renders q false; sufficiency is simply not enough.

So, the TRANSFERENCE principle is false. So we must conclude that the Consequence Argument is unsound.

Some remarks on the scope of my criticism: First, what I said is only a criticism of *one* version of the Consequence Argument. I think that similar criticisms apply to other versions van Inwagen defended, but for reasons of space I cannot discuss them here.[33] Second, what I said is *only* a criticism of the Consequence Argument; it is by no means an argument *for* compatibilism. Third, in the end I would conceive of the criticism more as a *challenge* for the argument, rather than as a rebuttal. As the argument was originally presented, it relied on a false transference principle (which, in effect, ignored the hyperintensionality of ground-theoretic locutions). So the crucial question is whether one can find a sound transference principle in accordance with a more general theory of ground.*

33 Another version of the argument employs the operator 'N', where 'Np' means 'p & no one has or ever had a choice about whether p' (see Van Inwagen (1983), ch. 3.10). An analysis of that locution can proceed in terms of rendering something true/false (see Finch and Warfield (1998)), and thus also involves the notion of ground (see Schnieder (2008) for an analysis of 'N' in terms of 'because'). Hausmann (forthcoming), drawing on my analysis, shows in a ground-theoretical framework how different versions of the Consequence Argument employing the 'N' operator fail (in particular, he also addresses van Inwagen's recent take on the issue from Van Inwagen (2015)).

* *Coda*. When I was a student, I had strong intuitions and quickly developed firm beliefs on many philosophical issues. One of those beliefs is partly responsible for this paper: As I described in the first footnote, my initial motivation to discuss Peter's Consequence Argument was my compatibilist conviction.

Unfortunately, I am not good at keeping things.

Over the years, I not only lost oh so many pens, books, scarfs, and the like; I also lost many of my philosophical convictions—including the belief in compatibilism. Even though I still haven't seen a version of the Consequence Argument that satisfied me, I somehow developed the strong feeling that the argument is correct in spirit; but I cannot really get my head around it. So I am

Bibliography

Bennett, J. (1984), *A Study of Spinoza's Ethics*, Indianapolis: Hackett.
Bolzano, B. (1813), [Beyträge] *Beyträge zu einer begründeteren Darstellung der Mathematik*, Prag: Caspar Widtmann.
Bolzano, B. (1837), *Wissenschaftslehre*, 4 vols., Sulzbach: Seidel.
Correia, F. (2010), "Grounding and Truth-Functions", in *Logique et Analyse*: 53, 251–279.
Correia, F. and Schnieder, B. (eds.) (2012a), *Metaphysical Grounding*, Cambridge: CUP.
Correia, F. and Schnieder, B. (2012b), "Grounding – An Opinionated Introduction", in Correia and Schnieder (2012a), 1–36.
Crusius, Ch. (1743), [De Usu] *De usu et limitibus principii rationis determinantis vulgo sufficientis*, Leipzig.
Finch, A. and Warfield, T. T. (1998), "The Mind Argument and Libertarianism", in *Mind*: 107, 515–528.
Fine, K. (2012), "Guide to Ground", in Correia and Schnieder (2012a), 37–80.
Hausmann, M. (forthcoming), "The Consequence Argument Ungrounded", in *Synthese*.
Kant, I. (1992), *Theoretical Philosophy 1755–1770*, edited and translated by D. Walford and R. Meerbote, Cambridge: CUP.
Kappes, Y. and Schnieder, B. (2016), "Anything at All – the Deepest and the Shallowest Question", in *Philosophisches Jahrbuch*: 123, 543–565.
Krämer, S. and Roski, S. (2017), "Difference-making Grounds", in *Philosophical Studies*: 174, 1191–1215.
Künne, W. (2003), *Conceptions of Truth*, Oxford: Clarendon Press.
Leibniz, G. W. (1989), [Essays] *Philosophical Essays*, edited and translated by R. Ariew and D. Garber, Indianapolis: Hackett.
Leibniz, G. W. (1686), [Discourse] "Discourse on Metaphysics", in Leibniz (1989), 35–68.
Leibniz, G. W. (1714a), [Monadology] "The Principles of Philosophy, or the Monadology", in Leibniz (1989), 213–225.
Leibniz, G. W. (1714b), [Principles] " Principles of Nature and Grace, Based on Reason", in Leibniz (1989), 206–213.
Lewis, D. (1981), "Are we Free to Break the Laws", in D. Lewis (1986), *Philosophical Papers II*, 291–297.
Nahmias, E., Morris, G., Nadelhoffer, T. and Turner, J. (2006), "Is Incompatibilism Intuitive?", in *Philosophy and Phenomenological Research*: 73, 28–53.
Raven, M. J. (2013), "Is Ground a Strict Partial Order?", in *American Philosophical Quarterly*: 50, 193–201.
Rescher, N. (2016), "Why is there Anything at all?", in *Philosophisches Jahrbuch*: 123, 220–234.
Rodriguez-Pereyra, G. (2005), "Why Truth-Makers", in *Truthmakers*: *The Contemporary Debate*, edited by H. Beebee J. Dodd, Oxford: Clarendon Press, 17–31.

eventually back where I once started as an undergraduate. Oh my. Thank you, Peter, for years of philosophical puzzlement.

Rodriguez-Pereyra, G. (manuscript), "Why Is There Something Rather Nothing? A Probabilistic Answer Examined".

Rosen, G. (2010), "Metaphysical Dependence: Grounding and Reduction", in *Modality*: *Metaphysics, Logic, and Epistemology*, edited by B. Hale and A. Hoffman, New York: OUP, 109–113.

Ross, J. F. (1969), *Philosophical Theology*, IN: Bobbs-Merrill.

Rowe, W. L. (1975), *The Cosmological Argument*, NJ: Princeton University Press.

Schaffer, J. (2012), "Grounding, Transitivity, and Contrastivity", in Correia and Schnieder (2012a), 122–138.

Schaffer, J. (2016), "Grounding in the Image of Causation", in *Philosophical Studies*: 173, 49–100.

Schnieder, B. (2004), "Compatibilism and the Notion of Rendering Something False", in *Philosophical Studies*: 117, 409–428.

Schnieder, B. (2008), "On What We Can Ensure", in *Synthese*: 162, 101–115.

Schnieder, B. (2010), "A Puzzle About 'Because'", in *Logique et Analyse*: 53, 317–343.

Schnieder, B. (2011), "A Logic for 'Because'", in *Review of Symbolic Logic*: 4, 445–465.

Schnieder, B. (2014), "Bolzano on Causation and Grounding", in *Journal of the History of Philosophy*: 52, 309–337.

Schnieder, B. (2016), "In Defence of a Logic for 'Because'", in *Journal of Applied Non-Classical Logic*: 26, 160–171.

Schnieder, B. and Steinberg, A. (2016), "Without Reason?", in *Pacific Philosophical Quarterly*: 97, 523–541.

Schopenhauer, A. (1864), [Wurzel] *Über die vierfache Wurzel des Satzes vom zureichenden Grunde*, 3rd ed., Leipzig.

Vallicella, W. F. (1997), "On an Insufficient Argument Against Sufficient Reason", in *Ratio*: 10, 76–81.

Van Inwagen, P. (1975), "The Incompatibility of Free Will and Determinism", in *Philosophical Studies*: 27, 185–199.

Van Inwagen, P. (1983), *An Essay on Free Will*, Oxford: Clarendon Press.

Van Inwagen, P. (1996), "Why Is There Anything At All?", in *Proceedings of the Aristotelian Society* (Suppl. Vol.): 70, 95–110.

Van Inwagen, P. (2009), *Metaphysics* (3rd ed.), Boulder: Westview Press.

Van Inwagen, P. (2015), "Some Thoughts on An Essay on Free Will", in *The Harvard Review of Philosophy*: 22, 16–30.

Uwe Meixner

Metaphysical Differences

Abstract: This paper addresses the disturbing phenomenon of fundamental, pervasive, and perennial dissent among the metaphysicians. As is shown in the paper, even some of the metaphysical opinions of the famous and very able metaphysician John Philosophus are far from compelling to other, no less able metaphysicians (and vice versa, of course). The apparently irreconcilable conflict of metaphysical opinions suggests the vanity of all metaphysics. However, this paper also points to a way in which metaphysics may nevertheless be worthwhile; it is, unfortunately, not the way in which John Philosophus does metaphysics.

It is an old complaint against metaphysics that the metaphysical views of the metaphysicians differ not only in less than fundamental respects but also *fundamentally*. They differ in fundamental respects – and by this I mean: they fundamentally contradict each other. There are more or less sophisticated arguments on both sides of every controversy in metaphysics, and most metaphysicians enjoy arguing, and enjoy winning an argument even more than arguing. This makes metaphysics seem vigorously alive. But to observers who manage to detach themselves at least for a while from their own stakes in metaphysics, the activities of the metaphysicians, whether in the past or in the present, can very well seem perfect illustrations of Ambrose Bierce's general definition of *philosophy*: "a route of many roads leading from nowhere to nothing."[1] If this were the truth of the matter, it would certainly be a tragic truth, since most metaphysicians – I take it – wish to travel the road to metaphysical truth, and not to *nothing*.

The endless clash of metaphysical opinions – usually without according intellectual justice to the opponent, always without a conclusive result once the dust has settled – can be frustrating and tiring. I hereby let you partake of my particular frustration. It's just a mood. I am not always in the same mood.

I will consider metaphysical differences – differences of metaphysical opinion – on some metaphysical distinctions, that is, on *metaphysical differences* in another sense of the expression "metaphysical differences." All controversies I will consider revolve around questions of the following form: Are all F (for example, all *beings*) G, or are some F not G? Are all F not G, or are some F G? All controversies I will consider are, in my view, intellectually unsatisfactory.

1 Bierce (1999), p. 144.

https://doi.org/10.1515/9783110664812-006

Consider the famous metaphysician John Philosophus. This name is an alias; the person I designate by it is, in fact, among us.[2] I find using the alias helpful for keeping my personal sympathy for that person at a safe distance from my somewhat harsh philosophical criticism of some opinions of his. They are his opinions, but I could as well attack the metaphysical opinions of other philosophers, and other philosophers could as well attack mine and in doing so have the same general purpose that I have here. John Philosophus is simply the paradigmatic highly distinguished living metaphysician at the center of my paper – in which I wish to make a certain general point about *metaphysics*.

1 Philosophus on existence

Now, one of Philosophus's metaphysical opinions is this: "The category 'thing' comprises everything there is, everything that exists (for I take a stern anti-Meinongian line about non-existents: non-existents simply don't exist: the number of them is 0)."[3] Many metaphysicians will applaud this view, others metaphysicians certainly won't. It is a minor – merely terminological – complaint against it that an ontological term which applies to *everything* – as the word "thing" does in the sense Philosophus accords to it – is not a category, but a so-called *transcendental*. The major issue is whether *everything there is* – in other words, *everything* – exists, as John Philosophus has it, or whether *not everything there is* – in other words, *not everything* – exists, as Alexius Meinong famously held. Now, I myself think that both philosophers are right: it all depends on whether you conceptually identify being and existence, as Philosophus does, or conceptually distinguish being and existence, as Meinong did. If being and existence are conceptually the same, then it is trivially true that everything there is exists – there is nothing particularly robust, staunch or *stern* in having this opinion (contrary to what Philosophus seems to think). If, however, being and existence are not the same conceptually, if existence means *actuality*, as Meinong believed, then it is false that everything there is exists, for it is simply false that everything there is is *actual*. For example, some states of affairs *are*, but are not actual; take the state of affairs that the moon is larger than the earth. Some properties *are*, but are not actual; take the property of being a flying unicorn. Even some individuals, or in

2 The original character of this paper is that of a speech (and as a speech it was originally presented). The paper's original character has been preserved in its printed version.

3 Van Inwagen (2007), p. 199.

another word: particulars, are, but are not actual; for example, someone who is not actual but might have been actual instead of me, someone who never was born and never will be born but might have been born to my parents instead of me.

Unfortunately, metaphysicians are prone to insist that their own opinions are right, that the opinions of opposing metaphysicians are wrong to the point of absurdity. They are not fond of irenic resolutions of their differences; they like to win. We cannot be quite sure in this respect about Alexius Meinong, who, after all, has been dead for quite a while, but I have a hunch that John Philosophus, even with the distinction between being and actuality in place, would still insist that Meinong was simply and absurdly wrong. If I may put some words into his mouth: "I have no idea what you mean by actuality as distinct from being. To the extent that I can understand the term 'actuality', it is synonymous with 'being'." Would it help if one pointed out that the state of affairs that the moon is larger than the earth *has being* – because it is (identical to) something – *but lacks actuality* – that is, *existence, in one of the legitimate senses of this word* – because it does not obtain? Would it help if one pointed out that the property of being a flying unicorn *has being* – because it is something – *but lacks actuality* – that is, *existence, in one of the legitimate senses of this word* – because it is not exemplified by anything actual? It might help, but I strongly doubt that it would help. As David Lewis once remarked, "Any competent philosopher who does not understand something will take care not to understand anything else whereby it might be explained."[4] And John Philosophus is certainly a very competent philosopher, indeed a formidably competent philosopher.

It is not unlikely that in reaction to the suggested modestly Meinongian attempts to convince him of the existence of the nonexistent, or to say it in a less Meinongian, *less playful manner*:[5] to convince him of the fact that some things do not exist – that in reaction to these attempts John Philosophus would sternly reject the idea that there is more than one legitimate sense to the word "existence"; there is just one sense, he is likely to insist, according to which sense existence is conceptually identical with being. And consequently (according to Philosophus), since *every thing is*, no thing does not exist, or in other words (accepting the all-

4 Lewis (1986), p. 203.

5 Meinong wrote: "Who is fond of paradoxical ways of expression could, therefore, very well say: objects exist of which it is true that such objects do not exist" (Meinong (1988), p. 9; my translation). The second part of this quotation (the part after the colon) is known as "Meinong's Shocker." One usually ignores the first part of the quotation where Meinong is clearly implying that (what came to be known as) "Meinong's Shocker" is merely a rhetorical, *playfully paradoxical way* of putting his central thesis.

encompassing sense Philosophus accords to the word "thing"): since *everything is*, nothing does not exist; and therefore, Meinong, who held that some things do not exist, is in error, absurdly in error.

I am not saying that to argue *like this* is wrong, but I do say that it is *arbitrary* and that it says nothing whatsoever against Meinong. Consider that many metaphysicians have believed that being itself is not univocal. Aristotle, for one, writes in Book IV of the *Metaphysics*, 1009 a (my translation): "For being is said in two senses, such that in the one sense something can come into being from non-being, but in the other sense it cannot." Yes, indeed, something can become *actual* from not being actual, but nothing can become *something* – that is, identical with something – from not being anything – that is, from not being identical with anything. Thus, Aristotle distinguishes two senses of being, one according to which *being* is *being actual*, another according to which *being* is *being something*. Believe me, Aristotle certainly believed that *some things are not actual*, and he certainly did not believe that some things are nothing; on the contrary, he believed that *every thing is something*. Now, Alexius Meinong identifies the meaning of the word "exist", which word (of Scholastic origin) was quite unknown to Aristotle, with *being actual*. Thus, Meinong says (or would have said in English), "Some things do not exist", or more dramatically, "Some things there are which do not exist", where Aristotle would have said, "Some things are not actual". Where, then, is the problem? Both philosophers hold what is logically the same, but Meinong alone, as we all know, gets the big howl, the derision, and the sneers – for no good reason.

John Philosophus, in turn, identifies the meaning of the word "exist" – which word, I repeat, was unknown to Aristotle – with *being something*. Thus, Philosophus says, "Every thing exists," or more dramatically, "Every thing there is exists", where Aristotle would have said, "Every thing is something". What Philosophus and Aristotle hold is logically the same, and what they hold is quite true, indeed, *trivially* true; but it says nothing whatsoever against Meinong; for Meinong was far from denying that every thing is something. If one asserted that *everything is actual* – well, then, indeed, one would be saying something against Meinong (and against Aristotle). But one would also be saying something which is false.

I am afraid John Philosophus will disagree. Perhaps Meinong and Philosophus will meet some day, and a higher authority will tell them who is right? In the meantime, I find the thought attractive that just like some people are actual who might not have been actual, so, by ontological symmetry, some people are not actual who might have been actual. They stay forever unactualized in the mind of God – who does not forget them and what they might have been and done if they had been actualized.

2 Philosophus on categories

I turn to another controversy. Many of John Philosophus's fellow materialists – notably, Jaegwon Kim and Donald Davidson – believe in *events*. There are also dualists who believe in events. I do. But John Philosophus does not believe in events, although he is somewhat tentative about this, as can be seen in the following quotation: "If, as I'd prefer to think, there are no events, if there are only substances and relations, then there is no thesis that can properly be called the identity theory."[6] Now, the conditional I just quoted is just false. Here is a thesis that can properly be called the thesis of the (token) identity-theory: "Every mental event is a physical event." This thesis obviously exists (both in Philosophus's and in Meinong's sense), no matter whether there are events or not. And if there were indeed no events, then this thesis would even turn out to be true, trivially true. What more can an identity theorist wish for? Many identity theorists will admit, though, that their thesis would turn out to be true not quite in the way they expect if there were no events.

However, what should really interest us here is, of course, something else: it is what Philosophus says he'd prefer to think: that there are no events. I don't know why anyone would prefer to think that there are no events. Ordinary language provides us with general and singular terms for events: "war" and "World War I," "assassination" and "the assassination of President Lincoln," "birth" and "the birth of Christ," "revolution" and "the French Revolution," "explosion" and "the supernova observed in 1054 A. D." We have idioms for characterizing events as actual or as non-actual: once an actual event is over, we say that it *happened*, or *took place*, or *ran its course*. If an event happened, it is an actual event. In contrast, the moon-landing of 1964 is an event that did not happen; it is a non-actual event, and so is the Coming of Christ in 2000 A. D. Note that we (or most of us) are quite ready to admit that there are non-actual events of the following stripe: events whose times – the times for them to happen – lie in the past and which simply did not happen.

John Philosophus would prefer to think that all of this sophisticated but entirely *ordinary* conceptual apparatus (which children master as soon as they can look forward to their *birthdays*) is good for nothing, since, as he says, he'd prefer to think that there are no events. What is such a negative attitude good for? It quite escapes me.

6 Van Inwagen (2007), p. 210.

Philosophus moves from being tentatively to being boldly negative when he says: "Even if there are mental events, I say, there are no pains."[7] *This* I have also heard from a fanatically devoted Wittgensteinian: from Peter Hacker. Whatever John and Peter *say*, there certainly *seem* (to everyone alive) *to be* pains, and they *seem to be* mental events of a special sort: subjective events. Why not say that there *are* pains and that they *are* subjective events? Other feelings besides pains are also subjective events. And sensations, dreams, hallucinations, episodes of thinking, episodes of strenuous willing, episodes of melancholy, episodes of listlessness, episodes of joy, episodes of religious emotion, temporal stretches of visual perception, temporal stretches of tactile perception, and so on – all of these are subjective events. Subjective events, insofar as they belong to *one* subject of experience, constitute the stream of consciousness of that subject of experience, its so-called *inner world* (or *inner life*). It is a very rich world, and for more than a century now not only psychologists but also philosophers have been willing (even eager) to explore it and describe it. The philosophical discipline devoted to this task is called "Phenomenology"; it was inaugurated by Franz Brentano and Edmund Husserl. Judging from his dismissal of pains, I am afraid John Philosophus will also dismiss all other subjective events, and therewith Phenomenology. If so, the only comment I have is that this is philosophically regrettable – regrettable from the *rational* point of view.

How could one convince someone who says that there are no pains that there are pains after all? It is very simple: pinch him where it hurts. But short of such drastic measures, which might land one in jail, the only method is to point out that our common discourse does strongly suggest that there are pains. Now, John Philosophus does not wish to be misunderstood: "I am not saying that if there are no events the eliminativists or the behaviorists are right. If there are no events, I contend, the mental is nevertheless real. For, even if there are no events, it is nevertheless true that some things think and have feelings. They really do have those properties. That they have those properties is as real and objective a feature of the world as anything is."[8] Well, I am glad to hear this. Philosophus's position seems to be that the event discourse about the mental – in particular, the event discourse about pain – can be wholly replaced by a substance-and-relation discourse about the mental (properties being counted as 1-term relations), without any significant loss. Perhaps the former discourse can indeed be wholly replaced by the latter, I am not sure; but even if this could be done, it would not mean that there are no pains, let alone that there are no mental events. For one thing, not mention-

7 Van Inwagen (2007), p. 210.

8 Van Inwagen (2007), p. 210.

ing something just does not mean that there is no such thing. For another thing, the event discourse about pain is just as likely to be able to wholly replace the substance-and-relation discourse about pain, without any significant loss, as it is likely to be the other way around. This point can be illustrated by a story:

> Ludwig – or Peter, or John – goes to the doctor. Sticking to his convictions not only in theory but also in practice, he is careful not to use the language of pain-events. He says to the doctor: "I am in pain with respect to my head and my stomach; moreover, I am in pain with respect to my left big toe." The doctor, who is without Ludwig's ontological scruples, jots down: "The patient complains about pains in the head, in the stomach, and in the left big toe." Then the doctor asks Ludwig: "How long have the stomach-ache, the headache and the pain in the big toe been going on? Do you have any idea what caused these pains? Did you already do something to make them go away?" Clearly, the doctor is using the language of pain-events – which language is anathema to Ludwig. Ludwig (or Peter, or John) pontificates: "I say, there are no pains." Imagine the amazement of the doctor.

This (fictitious) anecdote is *not* a fable for illustrating Bishop Berkeley's famous maxim: that we should think with the learned and speak with the vulgar. Its real message is that Ludwig is simply wrong (as are Peter and John): If Ludwig is in pain with respect to his head, his stomach and his left big toe, then there are, indeed, three pains – three (actual) pain-events – of which Ludwig is the subject: one in his head, one in his stomach, and one in his left big toe. And the converse is also true: If Ludwig is the subject of three pains: one in his head, one in his stomach and one in his left big toe, then Ludwig is in pain with respect to his head, his stomach, and his left big toe. The language of substances and relations in talking about pain is not privileged over the language of events; and vice versa, the language of events in talking about pain is not privileged over the language of substances and relations.

What does this linguistic fact suggest with respect to ontology? It strongly suggests that there are not only substances which are in pain but also pains, and, of course, it also suggests that there are not only pains but also substances which are in pain. Will this reconciliatory offer be accepted by John Philosophus, *or* by those other metaphysicians who, in diametrical opposition to John Philosophus, believe that there are *no substances*, hence no substances in pain; who believe that there are, as far as pain is concerned, only pain-events and their properties – of which the most important one is the property *being a pain* itself, with its species: being a headache, being a stomach-ache, being a toothache, being a toe-ache, and so on? My hopes for reconciliation among the metaphysicians are, in fact, infinitesimally small. For one thing, metaphysicians are not fond of reconciliation; they want to be victorious. For another thing – and this is a much more serious problem than the, perhaps narcissistic, opinionatedness philosophers so often display – the

majority of metaphysicians is enamored with monism: they have fallen for monism, they are fascinated by its alleged beauty; they covet the *one*-category ontology, they passionately desire existence to be univocal. However, the father of systematic metaphysics, Aristotle, was free from this infatuation with monism, was far from being a monist. "*To on pollachos legetai*" – "Being is said in many ways" is a dictum repeated over and over by Aristotle, and is repeated after him by Thomas Aquinas: "*Ens multipliciter dicitur*."

Willard van Orman Quine, in contrast, who has done much to reinstate metaphysics on the stage of contemporary thought, wants the landscape of ontology to be a desert landscape – not a Brazilian jungle, not an English garden (like the one in Munich), not a mountainous American, Polish or German woodland; Quine writes: "Wyman's [that is, Quine's philosophical dummy's] overpopulated universe is in many ways unlovely. It offends the aesthetic sense of us who have a taste for desert landscapes."[9] This is not humor, this, to my mind, is chilly superciliousness, and it leaves one with the eerie feeling that Quine's aesthetics is an aesthetics of death and dearth. Aren't we lucky, aren't we *blessed* that God is – also in this respect – not like Quine?

What about John Philosophus? Well, his universe looks even more lifeless than Quine's. John Philosophus's universe is a strange one in appearance. Look around you. What do you perceive? According to the metaphysical views of John Philosophus and many other philosophers, *you perceive nothing at all*. There are substances, of course, according to Philosophus, which have properties and stand in relations to one another, and which change, that is, lose old properties and gain new ones, move out of old relations with other substances and enter into new ones. But, according to Philosophus, all the properties and relations involved are abstract entities. The abstractness of properties gets special emphasis; Philosophus says: "It should be evident that properties, as I use the term, are as abstract as anything could be. They can in no way be 'constituents' (whatever that might mean) of concrete objects."[10] Abstractness of properties and relations means (among other things) that you cannot perceive them, not a single one of them; and this means that you cannot perceive that substances *have properties*, and *lose and gain properties*, or that substances *stand in relations*, and *move out of and into relations*. You cannot perceive this – because properties and relations are, according to John Philosophus, *abstract* entities. For illustration, note that there is *money in the concrete* – bills and coins – and *money in the abstract* – symbolized by certain figures on your bank statement. Money in the concrete you

9 Quine (2004), p. 179.

10 Van Inwagen (2007), p. 202.

can perceive, visually or otherwise, and you can perceive, visually or otherwise, that somebody *has* money, *gains* money, *loses* money in the concrete. But money in the abstract you cannot perceive, and you cannot perceive that somebody *has*, or *gains*, or *loses* money in the abstract; you can only infer, given appearances and your knowledge of the ways of the world, that a given person has, gains, or loses money in the abstract.

If assuming that properties and relations are abstract entities still does not seem to you to force you to conclude that all changes of substances in their properties and relations is unperceivable, just as all their having of properties and standing in relations is unperceivable, then you should recall that Philosophus does not believe in events and states of affairs. He declares:

> First, there is only one kind of concrete object: that which has traditionally been called 'substance' or 'individual thing'.[11] And there is only one type of abstract object. I will call this one type 'relation'. [...] Among relations there are 0-term relations, or propositions, 1-term relations (also called properties [...]), and 2-or-more-term relations, which I will call 'proper relations' (on the model of proper fractions and proper subsets). [12]

And John Philosophus adds:

> [S]ince, in my view, there are only substances and relations, there are no tropes or immanent universals. [...] [A]s far as I can see, the term 'trope' (as used by philosophers), and the term 'immanent universal' are perfectly meaningless. [13]

It always astonishes me how quick philosophers are to declare to be "perfectly meaningless" what other philosophers think, especially if it contradicts their own opinions. If metaphysicians resort to speaking like the inveterate enemies of all metaphysics – the logical empiricists – spoke, then this does not bode well for philosophy, and in particular for metaphysics. In any case, it should be clear from the quoted passages that John Philosophus has no proper place in his ontology for events and states of affairs, although, as we have seen, he can be somewhat tentative about events. Note that states of affairs are not propositions; "Propositions are things that have truth-values", says John Philosophus.[14] States of affairs do not

11 It must be noted that not even traditionally *substance* and *individual thing* have been the same category. For example, traditionally so-called *individual accidents* are traditionally not substances, though they are individuals and hence individual things (in Philosophus's wide sense of the word "thing").

12 Van Inwagen (2007), pp. 200–201.

13 Ibid., p. 202.

14 Ibid., p. 201.

have truth-values; they obtain and are facts, or do not obtain and are non-facts (or, as some Americans nowadays like to say, "alternative facts").[15] Now, every having of a property and standing in a relation is a state of affairs; every change in properties, every change in relations is an event. So, under Philosophus's assumption that there are no states of affairs and no events, there is nothing there for you to perceive when you look around you[16] – *except, perhaps, substances whose states and changes you cannot perceive*? But substances whose states and changes you cannot perceive are substances you cannot perceive. Thus, according to the metaphysical views of John Philosophus, when you look around you, *you perceive nothing at all*, as I said – which consequence constitutes a *reductio ad absurdum* of some of the metaphysical views of John Philosophus.

Will this impress John Philosophus? Not him, I'm afraid. If worse comes to worst, perception will be just another mystery for John Philosophus; he tells his readers: "I'm a metaphysician and am inured to mystery."[17] Perhaps he should better have said: "I'm a metaphysician and am inured to absurdity." He would be in illustrious company: a long series of great metaphysicians, beginning with

15 They are also *less fine-grained* than propositions: the *state of affairs* that John sits to the right of Peter is numerically identical to the *state of affairs* that Peter sits to the left of John; but the *proposition* that John sits to the right of Peter is numerically different from the *proposition* that Peter sits to the left of John. The same point can be made with respect to the state of affairs, respectively proposition, that ABC is an equiangular triangle and the state of affairs, respectively proposition, that ABC is an equilateral triangle (or the state of affairs, respectively proposition, that Harry loves Sally and the state of affairs, respectively proposition, that Sally is loved by Harry). Note that "that"-phrases are complex singular terms which are systematically ambiguous with respect to their object of reference: a certain proposition, *or* the state of affairs uniquely determined by that proposition. (And note: different propositions may determine the *same* state of affairs, as has just been amply illustrated.)

16 "I perceive that I have two legs" is a perfectly normal sentence, and a sentence I can verify directly by perception (of which I am conscious): I perceive (and am conscious of perceiving) the following *concrete* state of affairs: *my having two legs*, in other words, my having the *concrete* property *being two-legged* (or *having two legs*). But what does John Philosophus make of "I perceive that I have two legs"? He could treat "that I have two legs" as the logical object of "I perceive" and as designating *the proposition that I have two legs*; but this renders the sentence in question false, necessarily false; for propositions – being abstract objects – cannot be perceived by anyone. Philosophus must resort to an adverbial construal of "I perceive that I have two legs": "I am being I-two-leggy veridically appeared to" (and "I perceive that *he* has two legs" becomes "I am being *he*-two-leggy veridically appeared to"). This is certainly not what natural language and the phenomenology of intentional consciousness suggest. I am not saying that the adverbial construal of perception-sentences is wrong, but it seems to me *arbitrary*, an arbitrary, grotesque, and entirely unnecessary contortion.

17 Van Inwagen (2007), p. 201.

Heraclitus and Parmenides, including, near its (current) end, David Lewis and David Armstrong.[18]

The title of this conference is *Quo vadis, Metaphysics*? A serious question, indeed! A *First Philosophy* whose exponents flout the natural ontology of natural language, are madly in love with monism or unduly fascinated by revisionary radicalism; a *First Philosophy* whose exponents forever contradict each other, and forever argue one against the other, tacitly or not so tacitly intellectually contemning each other – where, in the words of David Lewis, "[o]ne man's reason is another man's *reductio*"[19] – *such a First Philosophy* is, it would seem to me, hardly *in* or *on* a good way. Where are you going, Metaphysics? From nothing to nowhere? From nowhere to nothing? I hope not.

3 Somewhat brighter perspectives

I must not allow this paper to end without suggesting a road away from "the highway to nothing," away from the mere clash of opinions where the, *for the moment*, strongest arguer (by dint of a quick intelligence and ready eloquence) will usually prevail – *for the moment*, but always not for long. I propose that we must, first of all, avoid gratuitous denials of entities (such as, the denial that there are subjective events, in particular, pains, or the denial that there are inanimate material objects). That there are arguments for such denials is, in itself, not enough rational justification for (seriously) proposing the denials (not as hypotheses to play with, but as something really to believe in), let alone for accepting them (in all seriousness): Where there are arguments, there are, or will be, counter-arguments, and where there is a smart arguer, there is, or will be, an equally smart or smarter counter-arguer. *What there is* is, it seems to me, not a matter of "who is the smartest of them all."

But *what should* – and, therefore, *could* – each one of us metaphysicians *do* other than argue – always unsuccessfully in the long run, no matter how smart we are – on the basis of his or her basic convictions (which, unfortunately, con-

18 What is absurd in the opinions of these great philosophers? Well, Heraclitus thought that everything moves, Parmenides thought that nothing moves, David Lewis thought that what *he can do* has something to do with what some counterpart of his does in some other possible world, and David Armstrong thought that he can have a perfect ontology of modality while being a perfect actualist. (For a detailed criticism of Armstrong's and Lewis's philosophies of modality, see my book *The Theory of Ontic Modalities*.)

19 Lewis (1986), p. 207.

tradict those of other metaphysicians) for his or her non-basic convictions (which, in their turn, contradict those of other metaphysicians)? This is what we should do: We should accept a common evidential basis and a common methodology. The evidential basis is ready at hand: natural language (including its many specializations, among them the language of science), which is the objective – hence intersubjectively available – mirror of human consciousness (which is the mirror of the world).[20] And the first methodological precept (besides the pledge of allegiance to logic) should be this: *Preserve as much of the prima facie ontology of natural language as is logically possible*.

Having said this much, I realize that it is a dream which is humanly impossible to realize – like Husserl's dream of philosophy "als strenge Wissenschaft [as a strict science]."[21] In *basic* matters, we humans tend to disagree fundamentally, even against the deliverances of natural language, which is, moreover, far from deciding everything that pertains to the *basic* matters. Having 2500 years of endless and heated philosophical conflict behind us, I have no hopes that this situation will ever change (as long as human history remains essentially *human* history). This means: I have no hopes that – external pressure (politics, religion, fashion) aside – *one* system of metaphysical propositions will ever be accepted by *all* metaphysicians, indeed, not even by the best metaphysicians.

What, therefore, should we metaphysicians do? The second-best thing to do – since the best is infeasible – is to accept *metaphysical pluralism*. Metaphysical pluralism is something else than metaphysical relativism, for under metaphysical pluralism the belief in objective metaphysical truth is being retained. In particular, every metaphysician is allowed to believe that his or her "system" tells the objective metaphysical truth (more or less completely), provided the system is logically consistent and has a sufficiently clear interpretation (whether a metaphysician is right in that belief is, of course, quite a different matter). Everything can stay as it is now, *with two significant exceptions*: (1) Inflammatory (eristic, polemical) language – like calling the opinions of others "perfectly meaningless" – must be strictly avoided. (2) The aim of debate must not be victory but finding, and stating a clearly as possible, the fundamental point of disagreement. Sometimes *no* such point will be found; then nothing stands in the way of fundamental agree-

20 What is called "ordinary language philosophy" is not the philosophy of ordinary language, not a philosophy of what ordinary speakers mean and ostensibly refer to when using ordinary language; it is an enterprise driven by anti-metaphysical ideology (which appears, on a closer look, not anti-metaphysical after all, but in the secret service of the metaphysics of scientistic naturalism).

21 See Husserl (2009), first published in 1911.

ment (about the particular point at issue). If, however, the fundamental point of disagreement *is* found, then all that remains is this: to agree to disagree, to agree to differ (in as clear a way as possible).

The effect of obeying these two imperatives will, indeed, not be the emergence of one unified consistent and comprehensive system of a pluralistic metaphysics – a system *e pluribus* (*categoriis et modis entium*) *unum*, a system into which all (professionally serious) metaphysicians invest their work (so that the system is *e pluribus unum* also in this other sense). *Such* a feat of human intellectual cooperation is beyond the pale of (merely) human possibility.[22] Rather, the effect will be the emergence of many unified consistent systems of metaphysics side by side, all of which are accepted by only some metaphysicians – sometimes by only one metaphysician. Some (many) of these systems will stand in logical conflict with one another; but since the nature of the conflict is clear in each case, and the conflict is without heat in each case,[23] this outcome – *metaphysical pluralism* – is the second-best outcome after the one outcome which is best but humanly unattainable (i.e., the all-encompassing and universally agreed on *e-pluribus-unum* system). The benefit to the general public – which, being human, craves metaphysics (sometimes more, sometimes less) – will still be great, since the general public is given – *not*, indeed, what the experts have agreed on is the metaphysical truth, but – what the experts have agreed on are crystal-clear and consistent fundamental alternatives in the quest for metaphysical truth (the existence of which truth is, *emphatically*, not denied). This is still helpful and a great gift of rationality, especially since it is accompanied by the following caveat: "We, collectively, do not know the metaphysical truth. But as individuals, some of us may indeed know it." Yes, indeed: some of us may know it. After all, whosoever is drawn into

22 Not only a universally accepted comprehensive system of presumed metaphysical truths is humanly infeasible, humanly infeasible is also a universally accepted comprehensive system of metaphysics which is (in the eyes of all metaphysicians) maximally coherent with the non-metaphysical aspects of presumed human knowledge. (This is easily seen: A comprehensive system of metaphysics should either include the proposition that God exists, or the proposition that God does not exist. But is the proposition that God exists more coherent with the non-metaphysical aspects of presumed human knowledge than its negation, or is it less coherent? This question, I submit, will forever – at least until the Last Day – stay moot among the metaphysicians.)

23 Obeying the two aforementioned imperatives will certainly result in there being less "fun" in metaphysical debates. But the kind of excitement that will be lacking (it can also be had from athletic contests or from chess games, as a participant in one way, as a spectator in another) is the kind of "fun" that we, as philosophers, should do without, at least in our professional lives. *We, as philosophers, are not and should not like to be gladiators, and we are not and should not like to be spectators of gladiators.*

the vortex of metaphysics, will, after a while (after reflection and getting older), not less firmly but ever more firmly uphold his or her metaphysical convictions, and these convictions may in fact – by chance, fate, or by the will of God – all be right, put together side by side: in conjunction; and many of them may, moreover, have been argued for in a rationally satisfactory manner (but note: it is rationally impossible to argue for every conviction one has). All of this may be the case, although there is not – nor, in all likelihood, ever will be – a corporate agreement among the metaphysicians to accept some particular metaphysical system. Thus, here they are: the disparate metaphysical convictions of the metaphysicians; here they are for us to accept a selection of them – if head and heart so dispose us – thoughtfully, coherently, clearly seeing the pros and cons. It is true: if metaphysics gets into contact with morality and religion (and this can hardly be avoided), the fire of polemical contest, which I wish to see extinguished in metaphysics, tends to flare up uncontrollably (sometimes leading not only to mental violence). But *metaphysical pluralism*, in the sense I described it, can be expected to soften even this rather unfortunate psychological effect – *if metaphysical pluralism is taken to heart*.

4 Appendix

(1) Castañeda on philosophical method

I am grateful to Francesco Orilia for drawing my attention to Hector-Neri Castañeda's book from the year 1980, *On Philosophical Method*. In some regards, I agree with what Castañeda says there; in other regards, I don't. Here is a regard in which I wholeheartedly agree with him (and how could one not agree with him in this regard?):

> In sum, *the philosophical given for each philosopher is the totality of his diverse experiences and the whole of each of his thinking idiolects with its syntactico-semantical contrasts*: experience and language united in an organic whole. [24]

This is true. But although we all live in one world and, as speakers of English, all speak one language, there does not appear to be a convergence between our experiences and thinking idiolects, at least none that is sufficient for establishing the philosophical – in particular, metaphysical – system we all agree upon (and would have every reason to believe true). On the contrary, philosophers dia-

24 Castañeda (1980), p. 47; italics in the original.

chronically and synchronically produce a plurality of incompatible metaphysical (hence philosophical) systems. And Castañeda is perfectly aware of this. He even thinks that this peculiar "state of the art" is as it should be and, in fact, could and should still be improved:

> The most crucial and urgent need of our time is *a plurality of comprehensive philosophical systems based on rich, complex, and abundant data all carefully analyzed*. Only a vital philosophical pluralism can prepare the way for dia-philosophy. [25]

Again, I wholeheartedly agree. However, Castañeda believes that an interesting *dia-philosophy* is (at least) possible; I do not believe that an interesting *dia-philosophy* is possible. What is *dia-philosophy*?

> The ultimate aim is the comparative study of maximal theories in order to establish, through isomorphisms among them, a system of invariancies [*sic*]. Such comparisons and the establishment of such isomorphisms and invariances is *dia-philosophy*. Naturally, we can at present have only dia-philosophical glimmers. The fully worked out systems of the future may, perhaps, give rise to systematic dia-philosophy. [26]

It is trivially true that differing systems will also have something in common. But if two systems not only differ, but are also *incompatible*, then it is likely that what they have in common is fairly trivial – therefore *uninteresting* – from the philosophical point of view. And the likelihood that the commonalities of two incompatible systems are philosophically trivial increases with the fundamentality of the disagreements between them. Unfortunately, the disagreements between incompatible metaphysical systems are very often (not always, I grant) *very fundamental*. What *metaphysically interesting* commonalities, for example, might there be between the metaphysical system of John Philosophus, for whom *actual substances* are the concrete basic objects and there are many basic kinds of abstract object, and the metaphysical system of another (also very real) metaphysician, for whom *actual and merely possible eventlike particulars* are the concrete basic objects and sets constitute the only basic kind of abstract object? It is certainly hard to see any metaphysically interesting commonalities between these two systems. Since it is unthinkable to exclude either one of the two systems (or both) from metaphysics or from philosophy, their very existence seems to demonstrate the impossibility of dia-philosophy. Philosophical pluralism and, in particular, metaphysical pluralism appear to be quite *unmitigable*.

25 Castañeda (1980), p. 20; italics in the original.
26 Castañeda (1980), p. 15; italics in the original.

(2) An argument of John Philosophus's

John Philosophus puts much stock in arguments (in fact, he can seem to argue *incessantly*). One should emulate him only to a certain extent, since, in truth, it is the propositions that ultimately matter in philosophy and, in particular, in metaphysics, not the arguments. After all, every argument starts with a proposition (usually, this initial proposition is a conjunction of propositions), and ends with a proposition.[27] No argument can make the proposition argued for – the conclusion of the argument – more believable than the argument's premise: the proposition (usually, conjunction of propositions) on the basis of which the conclusion is argued. It is, therefore, a very usual reaction to a logically sound argument simply to disbelieve its premise – as much as, or more than, one disbelieves its conclusion. And as long as metaphysicians do not commit themselves to an intersubjectively binding basis of metaphysical knowledge, *disbelieving the premise* will always be a perfectly rational countermove to any metaphysical argument whatsoever.

But here is an argument of John Philosophus's, an argument that founders already for purely logical reasons (so that *disbelieving the premise* is not necessary for disarming it, but certainly can be helpful for this purpose nonetheless); it is an argument that concerns "Cartesian unionism": the doctrine that the human person ("I") is a "union or amalgam or whole" of *mens* (or *anima*) and *corpus*:

> If Cartesian unionism is true, I am not the immaterial thing that Descartes calls my *mens* or *anima*. Suppose my body were annihilated and no new body replaced it. What would happen to me according to Cartesian unionism? Only one answer is possible: I should cease to exist, for, now that my body has been destroyed, there is no candidate for the office 'I' but my *mens*, or the *mens* that was formerly mine. And my *mens* can't be I, since it used *not* to be I – and, as we all know nowadays (I hope we all know this), if x is not identical with y, x is necessarily not identical with y. [28]

Regimenting the suppositions of this argument, we obtain:

> (Proposition 1) My body does not now exist, nor any replacement for it. (Let us suppose my body has been annihilated without replacement.)
> (Proposition 2) If I now exist and my body does not now exist, nor any replacement for it, then I am now my *mens*. (Given the antecedent of this conditional, there is now no other entity than my *mens* for me to be identical to.)

---

27 Indirect arguments constitute a special case. The premise of an indirect argument is the conjunction of its initial propositions *without its assumption for reductio*. The conclusion of such an argument is not its final proposition, the proposition with which it ends, but *the negation of its assumption for reductio*.

28 Van Inwagen (2007), p. 205; all italics are already in the original.

(Proposition 3) In the past, I was not my *mens*. (This is how Cartesian unionists – but certainly not only Cartesian unionists – would have it.)

Philosophus believes that the conjunction of these three propositions entails the following conclusion: *I do not now exist*.

Does this entailment really obtain? Philosophus seems to argue like this: Since in the past I was not my *mens* (according to proposition 3), I am necessarily not my *mens* (using the negative part of what one might term "Kripke's Law": that negative part is cited in full in the last quotation above[29]), and hence (as a logical consequence) I am now not my *mens*. Therefore (making use of proposition 2 [and elementary propositional logic]), I do not now exist, or my body or a replacement for it now exists. Therefore (making use of proposition 1 [and elementary propositional logic]), *I do not now exist*.

Philosophus seems to think that this result constitutes a *reductio ad absurdum* of Cartesian unionism *qua a form of dualism*, since, *as a form of dualism*, it must be taken to imply that the current non-existence of my body, even without a replacement for it, does not rule out my current existence.[30] Now, such a strong stipulation concerning the semantic content of "dualism" (that is, "psycho-physical dualism") can very well be made; I, for one, have no quarrel with it. However, the argument itself contains a serious logical problem: "I am necessarily not my *mens*" does not (does not *without further ado*) follow from "In the past, I was not my *mens*." It would indeed follow if both "I" and "my *mens*" were *rigid designators* across times and possible worlds (their being such designators would be a sufficient condition for their instantiating Kripke's Law, and hence for applying Kripke's Law); "I", as used by anyone, is indeed a rigid designator across times and possible worlds; but for "my *mens*" it is reasonably doubtful that it is a rigid designator across times and possible worlds.[31] If it is *not* such a designator (and it

29 The negative (second) part of *Kripke's Law*: If x is not identical with y, x is necessarily not identical with y. The positive (first) part of *Kripke's Law*: If x is identical with y, x is necessarily identical with y.

30 Note that not only Cartesian unionists, *qua dualists*, but also *materialists* are intended by Philosophus to find nothing objectionable in the three propositions the argument is based on, and quite a few of them, I expect, would take the argument to establish the following proposition: "If my body does not now exist, nor any replacement for it, then [with conditional necessity] I do not now exist" – which is just what *materialists* (including Philosophus) want, but *dualists*, of course, don't want.

31 Perhaps Descartes – who in fact did not believe that, in the last resort, *he himself* (this simple "I", not *the human being Descartes* as normally understood) is a union of *mens* and *corpus* – did indeed believe (in effect) that "my *mens*" (or "*mens mea*") is a rigid designator. But, as Philo-

may well be the case that it isn't), then concluding "I am necessarily not my *mens*" from "In the past, I was not my *mens*" is like concluding (falsely, of course) "Elisabeth is necessarily not my wife" from "In the past, Elisabeth was not my wife" (or like concluding "I am necessarily not Elisabeth's husband" from "In the past, I was not Elisabeth's husband").[32]

Weirdly, Philosophus thinks that Cartesian unionism is not an important position: "As far as I know, no one *is* a Cartesian unionist, and I don't propose to discuss at length a position no one holds."[33] Quite on the contrary, Cartesian unionism – the view that the human person is a "union or amalgam or whole" of *mens* (or *anima*) and *corpus* – is the most popular dualistic position, and of course, properly speaking, it isn't "Cartesian" at all; rather, it is "Aristotelian," or better still, it is simply *commonsensical*. The (pre-Christian) Roman poet Juvenal already exhorts us (whom he certainly takes to be unions of body and soul): "orandum est ut sit mens sana in corpore sano" (*Satire* 10, line 356). What has varied through the ages is the relative importance *unionists* (let's simply call them this way, omitting the misleading modifier "Cartesian") have accorded to the *body* in the mind-body union they believe in. Nowadays, the importance of the body for unionists is again at least as high as it was in Juvenal's time (whereas in the Christian Middle Ages it was very low).

Any unionist would be well advised to reply to Philosophus's above attempt to *absurdify* his or her position as follows: "In addition to my finding Philosophus's argument logically problematic, Proposition 2 is, in fact, not endorsed by me. I never am, nor ever can be, numerically identical to my *mens*, although, at times, I may very well coincide with my *mens*,[34] namely, if I should ever exist – I am inclined to think it possible – without my body, or any replacement for it, existing."

sophus intends the designation "Cartesian unionists," Cartesian unionists certainly need not be committed to all of Descartes's views in order to be "Cartesian unionists."

32 Neither "my wife" nor "Elisabeth's husband" are rigid designators across times and possible worlds. Obviously, both singular terms are *non-rigid*, both across times and possible worlds (in contrast to "I," as used by me or Philosophus, and "Elisabeth," as used by Philosophus or me).

33 Van Inwagen (2007), p. 206; italics in the original.

34 Remember Tibbles and Tib? Tib is (by definition) Tibbles without its tail. Tibbles and Tib are never ever numerically identical to each other, although at times (Tibbles having lost its tail) they *coincide*.

Bibliography

Aristotle (1989), *Aristoteles' Metaphysik*, Books I – VI, Greek and German, edited by H. Seidl, Hamburg: Meiner.

Bierce, A. (1999), *The Devil's Dictionary*, New York/Oxford: Oxford University Press.

Castañeda, H.-N. (1980), *On Philosophical Method*, Nous Publications.

Husserl, E. (2009), *Philosophie als strenge Wissenschaft*, Hamburg: Meiner.

Lewis, D. (1986), *On the Plurality of Worlds*, Oxford: Clarendon Press.

Meinong, A. (1988), *Über Gegenstandstheorie. Selbstdarstellung*, Hamburg: Meiner.

Meixner, U. (2006), *The Theory of Ontic Modalities*, Heusenstamm: Ontos.

Quine, W. V. O. (2004), "On What There Is," in: W. V. O. Quine, *Quintessence. Basic Readings from the Philosophy of W. V. Quine*, edited by R. F. Gibson, Cambridge, MA: Belknap Press, 177–192.

Van Inwagen, P. (2007), "A Materialist Ontology of the Human Person," in: *Persons: Human and Divine*, edited by P. van Inwagen and D. Zimmerman, Oxford: Oxford University Press, 199–215.

Anna-Sofia Maurin

A Van Inwagenian Defense of Constitutionalism

Abstract: According to Peter van Inwagen, the claim that metaphysics is in the business of *explaining* the nature and existence of whatever inhabits mind-independent reality *makes no sense*. Likewise, according to van Inwagen, constitutionalism – the view that bundles of properties constitute charactered objects – is 'not even false' and concomitant notions like those of a 'constituent', an 'immanent universal', and a 'trope' are incomprehensible. In this paper, I investigate both of those claims. I argue that the first – that it makes no sense to understand metaphysics as in the business of explaining – makes sense (!) only if understood as a corollary to the claim that it makes no sense to say that the relations that (hierarchically) structure mind-independent reality are *inherently* explanatory. This latter claim *is* problematic, and could be what proponents of the explanatory approach have in mind when they claim that metaphysics is in the business of explaining (although it certainly doesn't have to be). Suppose it is, and suppose the claim turns out to be unacceptable, perhaps even nonsensical. Then van Inwagen's first claim is true. But this does not mean that we have to accept his second claim. For, constitutionalism can be formulated *without* the assumption that metaphysics – and hence the relations which structure reality – is explanatory. Nor does Quineanism, even in its distinctively van Inwagean variety, require that properties be understood as existing 'apart' from the objects they characterize. Therefore, it makes sense to say (although it may be false) that properties 'make up' objects, that properties exist 'in' the objects they constitute, and that properties are immanent universals or tropes.

1 Introduction

In a recent paper in a volume dedicated to the philosophy of Peter van Inwagen, Michael J. Loux complains that, although van Inwagen emphatically *denies* that saying of an object that it exemplifies, say, greenness, counts as a substantive explanation of the fact it is green, he nevertheless takes *character to be grounded in properties*: a view "most of those involved in the debate would take … to involve a genuine explanation" (Loux (2017), p. 15, *fn*. 16). Van Inwagen does not agree. True, most of those involved in the debate *would* take saying that the object exemplifies greenness to genuinely explain why the object is green. They are, however, "*wrong, wrong, wrong*" (Van Inwagen (2017), p. 346). To say of one or several particular(s) that it (or they) is (or are) green means that there is one property said of it (or both). Yet there is no explanation here.

https://doi.org/10.1515/9783110664812-007

Why does Loux think van Inwagen is (reluctantly) doing something he officially (and emphatically) denies? Why does van Inwagen deny doing what many (most?) philosophers take to be both unavoidable and unproblematic? In this paper, I approach these and related issues by trying to prize apart the ontological project van Inwagen is interested in, from the sort of (explanatory) endeavor Loux and others engage in. Because those projects are often couched in very similar (identical even!) terminology, it is perhaps not surprising that, although, as I will try to demonstrate, van Inwagen's is *not* an explanatory project, it may nevertheless seem like it is. Seeing that it isn't lets van Inwagen off the Louxian hook. But, as I will argue next, it lands him on another. For, given a van Inwagean understanding of ontology, his critique of the view that, what exists when an object is *F* is something that is (partly) constituted by *F*-ness (henceforth: *constitutionalism*[1]), and of various ideas closely related to that view (including that of a 'trope', an 'immanent universal' and a 'constituent'), misses its mark. The upshot is this: Although his mistake is understandable, Loux is wrong in accusing van Inwagen of doing what he claims not to do (namely explain). But van Inwagen is *also* wrong when accusing constitutionalism (or, constitutionalism$_+$ as I will be calling the 'van Inwagean' version of the view) of being "not even false" (Van Inwagen (2011), p. 396). And both are arguably wrong for the same reason. They both judge an account of what exists (an ontology), without taking into consideration in response to which 'task' the view is being formulated. The moral of this paper – if there is one – is simple. Don't.

2 The problem of character

Look around you. A number of things seem *prima facie* evident. There are objects (concrete particulars), both (very) big, medium-sized and (very) small. Objects are 'charactered', meaning that they have a (complex) nature. Because they have the natures they do, objects resemble one another more or less closely. Because the natures they have is complex, two objects may resemble one another in one respect, yet be very different overall. According to D. M. Armstrong, these are all *Moorean* facts, by which he means that they are facts "even philosophers should not deny" (Armstrong (1997/1980), p. 102).[2] They are facts, moreover, such that,

1 Not to be conflated with Lockean constitutionalism!

2 According to Armstrong, more precisely, a Moorean fact is "one of the many facts that even philosophers should not deny, whatever philosophical analysis they give of such facts" (Armstrong (1997/1980), p. 102). And according to Lewis (1996), p. 549, it is "one of those things that

although they should be accepted, any comprehensive philosophy *must explain why they obtain*.[3] In this, Moorean facts "constitute the compulsory questions in the philosophical examination paper" (ibid.). Consider one of these facts – the fact that there are charactered objects – and consider the concomitant task of accounting for why that fact obtains (in order to solve 'the problem of character'). According to Loux – but certainly not *only* according to Loux – there are three possible solutions to this problem:[4]

> *Primitivism*: objects have the character they do because they have the character they do (it's primitive!);
> *Relationism*: objects have the character they do in virtue of exemplifying/partaking in some separately existing universal or Form; and
> *Constitutionalism*: objects have the character they do in virtue of the properties that constitute them.

These solutions differ, among other things, in how many categories they posit. Primitivism, first, is a one-category theory. Relationism and constitutionalism, on the other hand, both explain the nature of objects with reference to something – properties – that is *not* (or at least not obviously) an object. In this sense these are both at least two-category ontologies.[5] Of course, whether an object's constituents and the object they collectively make up, are 'ontologically on a par', or whether the fact that the constituents *ground* the existence and nature of the object means that only the former *really* exist (whatever that may mean), is an issue on which we may disagree. A majority of the constitutionalists who are also bundle-theorists (and hence hold that what makes up the object are – bundles of – properties), tend to think that theirs is a one-category theory according to which only properties (fundamentally or 'really') exist. Exactly how many categories constitutionalists who are also substrate-attribute theorists commit to is less clear. On e.g., Armstrong's view (Armstrong (1978, 1997)), although a (thick) object exists and has the character it does *because* it is made up from a thin particular (a substrate) in which a number of (universal) properties are instantiated, once the thin particular exemplifies the universals, there exists also a state of affairs. On

---

we know better than we know the premises of any philosophical argument to the contrary". *Cf.* also e.g., Nolan (2005), pp. 208f.

3 Where this explanation should not refer to the *cause* of the obtaining of the fact but, rather, to its (non-causal/metaphysical) 'grounds'.

4 Similar-sounding classifications of the solutions to the problem of character can be found in e.g., Wolterstorff (1970, 1991) who, in turn, attributes this way of classifying things to Bergmann (1967).

5 But cf. Paul (2002, 2017).

at least one interpretation of Armstrongian constitutionalism, therefore, when a charactered object exists, at least three (possibly four) sorts of things exist (thin particulars, universals, states of affairs (and objects)).[6] Common to all versions of constitutionalism is in any case the idea that *objects have an ontological structure*. Common to primitivism and relationism, on the other hand, is the idea that they don't.

3 Explanatory failure

Whether primitivism, relationism and constitutionalism constitute acceptable responses to the problem of character depends on whether or not one thinks that those responses count as (decent enough) explanations of the relevant facts. That they do can however be doubted.

Whether primitivism constitutes an acceptable response to the problem of character, first, depends on whether or not saying that, that there are charactered objects is primitive, counts as an explanation of that fact. According to Armstrong it does not. For, he claims, to account for the fact that objects are charactered – or that several objects have the same character – by saying 'it's primitive!', is to be like "an Oxford Philosopher of yore", who (Armstrong (1997/1980), p. 103):

> ... keeps on saying that he does not deny that many different objects are all of them red, but what this ostensible sameness is he refuses to explain (except to say that it is ultimate and irreducible).

According to e.g., Lewis (1983), however, it does. For, although a systematic philosophy "must indeed give an account of any purported fact", there is more than one way of doing this (Lewis (1983), p. 352):

> Not every account is an analysis! A system that takes certain Moorean facts as primitive, as unanalysed, cannot be accused of failing to make a place for them. It neither shirks the compulsory question nor answers it by denial. It does give an account.

What Armstrong's criticism – as well as Lewis' defense – illustrates, is just how important exactly how we understand the *nature* of explanation is to our evaluation of any candidate solution to the problem of character. For primitivism to pass muster, we have just seen, we must allow there to be explanation even in

6 To make presentation easier, I will in what follows talk as if constitutionalism equals constitutionalism in its bundle-theoretical guise.

cases where the explanans arguably doesn't 'add' anything to the explanandum (other than that 'there is nothing more to add', that is). Or, if you like, it must be such that explanation – at least in the limiting case – can be reflexive: that some things so-to-speak 'self-explain'. For primitivism to qualify, in other words, we must adopt what is arguably a rather unorthodox understanding of explanation.

Suppose we find this acceptable. Then primitivism is a possible (albeit arguably a 'last resort') response to the problem of character. Is the same true of its rivals, relationism and constitutionalism? *Prima facie*, yes! According to relationism, first, objects have the character they do *because* or i*n virtue of* their partaking in some distinct (and categorically different) Form or property. In Plato's words (*Phaedo* 100c3-5):

> I think that, if there is anything beautiful besides the Beautiful itself, it is beautiful for no other reason than that it shares in that Beautiful, and I say so with everything.

This is arguably explanation in a much more substantial and, not least, familiar sense than that to which the primitivist must appeal for her response to the problem of character to go through. The same seems to be true also of constitutionalism. Here the explanation is in terms of bundles of properties, but still the explanans is distinct from the explanandum, and explanation is hence informative and irreflexive.

Constitutionalism has however been accused of ending up in explanatory failure of another sort. The existence and nature of bundles of properties fail to explain the existence and nature of charactered objects, is the idea, because the former are so different from the latter, that how one gives rise to, and hence 'enlightens' us about, the other is a mystery. In this, constitutionalism is more like emergence than it is like explanation. Given the base (a bundle of properties), what emerges (the charactered object) is 'surprising'.

What is the source of this (supposed) explanatory failure? Hardly, or at least not only, the fact that the explanation in question involves a 'traversing of category lines'. Case in point: relationism. The relationist explains charactered objects with reference to the distinct and categorically different properties in which they partake. Or, perhaps better put: *the character of objects* is explained in that way. For, according to relationism, what the existence and nature of the Platonic Form explains is not the existence, but only the nature of the concrete object. And learning about the nature of the properties in which that object partakes *should* lead us to expect something about the nature of the object itself. Given the property, the object's character is *not* surprising. But, then, although relationism explains something about something belonging to one category (concrete object) with reference to something belonging to another very different category (abstract

property), this does not in and of itself mean that the view ends up in explanatory failure. Whether the difference between explanandum and explanans is great enough to threaten an explanation's status *as an explanation* ought hence to be judged, not just on the basis of what we take the categorical home of explanandum and explanans to be, but also on the basis of *what about the explanandum* the explanans is supposed to enlighten us. This is why relationism probably doesn't fall prey to this sort of objection, but constitutionalism does. For, on the latter view, the existence and nature of some properties does not only explain the nature of some object. It also makes or grounds or explains the existence of that object. Exactly how that happens is then what the critic finds mysterious.

4 The problem of character$_+$

At first glance, van Inwagen and Loux appear to be discussing the same problem, as well as offering and then evaluating the same sorts of solutions to that problem. Van Inwagen too distinguishes between what I above call primitivism, relationism and constitutionalism. He too finds constitutionalism problematic. And he does so for what superficially looks like the same (explanatory) reason as that discussed in the previous section. First impressions are however wrong. Van Inwagen and Loux are *not* discussing the same problem, which means that, although named the same, van Inwagean primitivism, relationism and constitutionalism (henceforth: primitivism$_+$, relationism$_+$ and constitutionalism$_+$ to distinguish these views from their similar-sounding cousins) are not solutions to the problem of character. If anything, they are solutions to the problem of character$_+$ (but even that is something of a misnomer, as we shall see in a moment).

What spells the difference is, once again, *explanation*. As we have just seen, the problem of character is that of explaining why the Moorean fact that there are charactered objects obtains. Yet, according to van Inwagen, *that* there are charactered objects (and whatever else there may turn out to be), rather than the starting point, is the end-result of 'analysis'. Van Inwagen is a Quinean. He looks to our best theory (or theories) and attempts to determine (using Quinean methods) what sorts of entities they commit us to. Rather than asking for an explanation of *why* there are the things there are (and expecting an answer in terms of what grounds or makes up the things in question), therefore, van Inwagen asks *what* (*kinds of*) *things there are* (expecting nothing more nor less than a list of kinds of things in response). *His* 'problem of character' is hence not so much a problem as it is a task: that of listing our ontological commitments. Just as before, there are (at least) three possibilities: primitivism$_+$, relationism$_+$ and constitutionalism$_+$:

Primitivism$_+$: our best theory commits us to the existence of (ontologically unstructured) charactered objects;
Relationism$_+$: our best theory commits us to the existence of (ontologically unstructured) charactered objects and abstract universals or Forms; and
Constitutionalism$_+$: our best theory commits us to the existence of (ontologically structured) objects made up from (universal or particular) properties.

Van Inwagen is a self-professed (albeit reluctant[7]) relationist$_+$. According to him, there are some sentences of our best theory which commit us to the existence of, not just objects, but properties as well. Properties are universal and abstract. Hence, van Inwagen thinks they exist necessarily, can be actually and even necessarily uninstantiated, are abundant, causally inert, imperceptible, and do not exist in space and time. Therefore, properties are distinct from the objects that have them. Which means that objects are ontologically unstructured. Hence, relationism$_+$ is true.

Primitivism$_+$, relationism$_+$ and constitutionalism$_+$ have many things in common with primitivism, relationism and constitutionalism. Primitivism$_+$ posits the same number of categories as its counterpart (one!), and the two views take the same stance on the ontological structure of charactered objects (they don't have one!). The same goes for relationism$_+$ and relationism, and constitutionalism$_+$ and constitutionalism, respectively. Yet, neither primitivism$_+$, relationism$_+$ or constitutionalism$_+$ *explain*. Hence, although all of primitivism, relationism and constitutionalism can be criticized for failing to explain, or for not explaining enough (*cf*. Sect. 3), the same is not true of their $_+$-marked cousins.

Why don't primitivism$_+$, relationism$_+$ and constitutionalism$_+$ explain? The short answer, according to van Inwagen, is that they don't explain because there is nothing there *to* explain. "[N]o set of statements among all possible sets of statements counts as an explanation of what it is for a particular to have a property or for two distinct particulars to have the same property" (Van Inwagen (2011), p. 398; cf. also Van Inwagen (2017), p. 347). Pointing this out is one part of van Inwagen's more general attack on 'explanation-driven' metaphysics. In giving the impression that one is proceeding somewhat as science does (producing explanation, comparing and weighing explanation, etc.), metaphysics trades on the respectability of science. But if there is nothing there to explain, this respectability is only illusory, precisely because the supposed similarity with science (which *does* explain) is mere chimera.

7 He would prefer the desert landscapes promoted by the primitivist, but feels compelled to go where his methodology takes him. *Cf*. the first section of Van Inwagen (2004), which is called 'it would be better not to believe in abstract objects if we could get away with it'.

One way to understand the claim that there is nothing to explain is simply as the Quinean claim (set out above) that the task of metaphysics – or, perhaps better, of ontology – consists in listing our ontological commitments, *nothing else*. Listing is not the same as explaining. Hence, there is nothing to explain. But van Inwagen often seems to be saying something more or stronger than this. He seems to be saying that the alternative task, promoted by the likes of Loux, of explaining why the things we are committed to exist and are the way they are (a task that does not obviously compete with, but could be taken to complement, the van Inwagean one), is *impossible* or even *nonsensical*. Why is that?

5 Ontological structure and explanation

One possibility is that van Inwagen thinks talk of metaphysical explanation makes no sense, because talk of ontological structure, talk of constituents, and, in general, talk of reality as somehow 'layered', makes no sense (Van Inwagen (2014a)). That he doesn't think *those* things make any sense is *very* clear (Van Inwagen (2011), p. 393).

> I do not understand the idea of ontological structure or, indeed, any of the ideas with which one finds it entwined in the various constituent ontologies. I do not understand the words and phrases that are the typical items of the core vocabulary of any given constituent ontology. 'Immanent universal', 'trope', 'exist wholly in', 'wholly present wherever it is instantiated,' 'constituent of' (said of a universal and a particular in that order): these are all mysteries to me.

That van Inwagen assumes that ideas about ontological structure must go hand in hand with (what he takes to be very objectionable) ideas about explanation, can be gleaned from his criticism of constitutionalism$_+$ (which, in his view, is really constitutionalism in sheep's clothing), at least in one interpretation. To van Inwagen, constitutionalism makes no sense, because (Van Inwagen (2011), p. 392):

> It ought to be as evident that there is no sense of 'constituent' in which unsaturated assertibles [van Inwagen's term for properties] are constituents of concrete particulars as it is that there is no sense of 'extraction' in which a physical tool [a 'forceps'] can be of use in the extraction of a cube root.

There are at least three ways to interpret this worry. On one interpretation, it is a worry about what properties – properly conceived – can and cannot do. On another, it is a worry analogous to that raised against constitutionalism given an explanatory approach to metaphysics, according to which a concrete particular can-

not be 'extracted' from a bunch of properties because this would be 'surprising' and hence constitute an explanatory failure. And on a third, it is a general worry about the (nonsensical) sense of explanation talk of ontological structure invokes. The first interpretation will be discussed in the next section. The second interpretation, although natural, makes no sense from a van Inwagean point of view. If there is nothing there to explain, if talk of metaphysical explanation doesn't even make any sense, there is hardly any interesting distinction to be made between good and bad or successful and unsuccessful (nonexistent) explanations. This leaves us with the third interpretation. On this view, what is problematic about constitutionalism and constitutionalism$_+$ (as well as about adopting an explanatory approach to metaphysics generally) is the *inherently explanatory* ontological structure these views commit us to.

That a commitment to metaphysical explanation is inextricably bound up with a commitment to ontological structure is a commonly held view. In fact, this is arguably the core-thesis of so-called 'grounding theory' (*cf*. e.g., Audi (2012); Fine (2012); deRosset (2013); Raven (2015)). Van Inwagen is rightly suspicious of this way of understanding whatever worldly relations structure reality. Explanation, one might argue, is inherently epistemic. But then, if constitution or grounding or whatever other worldly relation one takes to be primarily in charge of structuring reality is objective and mind-independent, what could saying of *it* that it is explanatory even mean? At best, the sense of being explanatory characteristic of worldly grounding/constitution is very different from explanation in the ordinary sense. If so, what sort of information is calling it 'explanatory' supposed to convey?

For the purposes of the argument set out here, suppose, therefore, that *inherently* explanatory yet mind-independent structuring relations make no sense. Does this mean that constitutionalism (and constitutionalism$_+$) must be rejected? Only if we assume that ontological structuring relations, if they exist, must be inherently explanatory. There is however reason to reject this assumption. As we have seen, a common view among proponents of the explanatory view – let's call them 'grounding theorists' – is that, when the more fundamental grounds or constitutes the less fundamental, the more fundamental also *explains* the less fundamental. This claim is hard to make sense of if explanation is as we normally take it to be – pragmatically constrained, subjective, epistemic. For then, if grounding is explanatory – if it is explanatory by *nature* – what might this mean other than that it, too, is pragmatically constrained, subjective, and epistemic. Yet grounding, according to a majority of the grounding theorists (for an exception, *cf*. e.g., Dasgupta (2017)) is a worldly, mind-independent relation. Now, not all grounding theorists take grounding to be explanatory *by nature*. In fact, according to most grounding theorists, although grounding *is* explanatory in the sense that if *a*

grounds *b*, the nature and/or existence of *b* is explained by the nature and/or existence of *a*, this is not because there is explanation *in nature*, so to speak. Or, it is, but only in the sense that the (pragmatically constrained, subjective, epistemic) explanations we give sometimes track (distinct) worldly, objective grounding relations (somewhat like most people take causal explanations to track (distinct) causal relations). On this view – so-called 'separatism' – explanation is as we normally take it to be, yet certain types of worldly relations are understood as playing a special role in (some of) them.

But if the worldly structuring relations are distinct from the explanations they back, it makes sense to say that those relations could obtain *without* backing any explanation at all. If explanation is essentially subjective, one obvious such case would be a world with no minds. If explanation is essentially mind-dependent, this would be a world without explanation, yet the world could still be hierarchically structured. In fact, there is no reason why we should not allow for this possibility even in our own – mindful – world. We could imagine all sorts of relations obtaining, but as long as there is no one there to acknowledge them, there is no explanation. And if you are a grounding skeptic of the kind who thinks that, although all sorts of relations play a role in all sorts of explanation, there is no reason to think that any *special* kind of – 'explanatory' – relation does, this phenomenon could be pervasive. To be 'explanatory' on this view, a relation must back an explanation, but doing so is not inherent to the relation or to the explanation, but is dependent on various extraneous facts in the situation at hand. In fact, if you are a grounding-skeptic of a maximally strong variety, if you are a van Inwagean skeptic, you don't think *any* explanation *ever* tracks *any* non-causal relation. Yet, it ought still to make sense to say that those worldly relations obtain. Just as it makes sense to say that they sometimes obtain without backing an explanation. But then, the obtaining of ontological structuring relations can be distinguished from their obtaining *and backing* (separate) explanation. Which means that, although the van Inwagean objection to constitutionalism goes through (given that constitutionalism is understood as the view that ontological structure *is* explanatory structure), the (same) objection to constitutionalism$_+$ does not. Van Inwagean constitutionalism$_+$ is the idea that the world is ontologically structured. But this idea is here distinguished from the idea that, if *a* ontologically 'makes up' *b*, *a* also explains *b*. Whatever ails constitutionalism$_+$, then, it cannot be that the structures it invokes *must* be explanatory.

6 Van Inwagen on the nature of properties

Suppose that there is a sense of ontological structure, in which some things (non-causally) depend for their existence and/or nature on other things. And suppose that constitutionalism$_+$ is the view that charactered objects have ontological structure in this sense. Does this mean that van Inwagen's critique of that view misses its mark? Not necessarily. For on one final interpretation of van Inwagen's objection to constitutionalism (or, constitutionalism$_+$, as we are now focusing on that view), the problem is that, although there is a respectable sense in which some things depend on, by being made up from, other (kinds of) things, charactered objects cannot be made up from properties *because properties are not the sort of thing that can make up objects*.

According to van Inwagen, as we have seen, there are properties. The reason for this, more precisely, is that there are truths that commits us to properties; there are truths for which no property-free paraphrase exists. If e.g., <Spiders and insects share many important anatomical characteristics> is true, it must be because there exists, besides spiders and insects, anatomical characteristics. Anatomical characteristics are *unlike* spiders and insects. They are *not* objects; hence they are properties. According to Van Inwagen, properties are abstract. As such, they are *very* different from the sorts of things with which we are normally acquainted. Accepting them, he thinks, means having to accept – on top of the sorts of things that are (comparatively speaking) well studied and well understood by our best sciences – something non-spatiotemporal, acausal and non-perceivable. Properties, thus understood, must exist apart from the objects they characterize. Hence, it does not make any sense to say that properties 'make up' or constitute charactered objects.

But why does van Inwagen think that properties must be non-spatiotemporal, acausal and non-perceivable? According to Quine – or at least according to the Quinean – the results of subjecting our best theory to Quinean quantificational analysis is a "large class of one-place open sentences that you believe are satisfied" (Van Inwagen (2011), p. 399; *cf*. also Quine (1948)). Listing the 'satisfiers' means listing 'what there is', but there is not – or at least not obviously – anything in this list that tells us anything specific about the nature of the entities it mentions. Van Inwagen is however not a 'pure' Quinean. InVan Inwagen (2014b), Ch. 9, he distinguishes between (1) ontology understood as the study of the ontological structure of objects (an explanatory – and hence according to van Inwagen objectionable – approach to ontology); (2) ontology understood as concerned with answering the ontological question: What is there? (what van Inwagen calls 'the bare Quinean conception'), and; (3) ontology understood as concerned with

answering the ontological question *in terms of a specification of the ontological categories*. Van Inwagen prefers to do ontology in the sense of (3). This is because he thinks whatever ends up on the list of existents will belong to this or that natural class, where these natural classes make up the ontological categories.[8] Which means that, on top of his commitment to a Quinean methodology, van Inwagen needs to accept and adopt a theory of the categorical nature of those things we find ourselves ontologically committed to. More precisely, van Inwagen needs to accept and adopt a *theory of properties*.

According to constitutionalism$_+$, properties exist in the objects they 'make up'. Despite van Inwagen's protestations, understanding properties in this way is *not* unusual or in general considered very strange or unfortunate. According to so-called immanent realists and trope theorists alike, since properties are spatiotemporal, they are causally efficacious (according to some, they are *the* primary bearers of causal power) and perceptible. Rather than the weird exception, this way of understanding the nature of properties is often defended with reference to its superior 'scientific' respectability. Instead of taking properties to be unobservable, causally impotent, and in general strangely ethereal, this way of understanding properties brings them down to earth and makes it possible for science to directly investigate them. All of which are virtues van Inwagen arguably appreciates. So, why won't van Inwagen consider understanding the nature of properties in this way?

Van Inwagen admittedly doesn't have a lot to say on this topic. In Van Inwagen (2004), p. 131, he describes his own understanding of properties as true, but nearly vacuous. It is a way of "specifying the property role" that is independent of, and slightly more informative than, "specifying this role via the apparent quantifications over properties that are found in our discourse." It does seem like he takes the features he attributes to properties to be features properties have in virtue of being abstract, though. And he does have a few things to say about what it means to say of something that it is abstract. His point of departure is this (Van Inwagen (2004), pp. 108–109):

> …it is possible to divide the terms and predicates we use in everyday and scientific and philosophical discourse into two exhaustive and exclusive classes by a very simple method. We stipulate that one class shall contain the terms and predicates in the following list: 'table', 'the copy of War and Peace on the table', 'Mont Blanc', 'the Eiffel Tower', 'Catherine the Great', 'neutron star', 'intelligent Martian', 'elf', 'angel', 'god', and 'God'. We stipulate

8 If the classes are 'large' and 'high', they constitute the primary categories, if they are natural proper subclasses of some large and high natural class, they constitute the *secondary*, *tertiary*, etc. ontological categories (*cf.* Van Inwagen (2014b)).

> that the other shall contain 'number', 'the ratio of 1 to 0', 'proposition', 'sentence'… 'property', 'angle'…, 'possibility'…[etc.] We then ask philosophers … to place each term of our discourse … in the class where it will be most at home.

Following this procedure, van Inwagen claims, "will yield pretty consistent results". Although, he admits, because some of the terms on the list of paradigms are ambiguous they will be differently understood (and hence classified) by different philosophers, and because "some philosophers may have idiosyncratic theories about the items in the extensions of some of these terms" those too might diverge from the 'norm' (ibid.). The terms in the first box pick out the *concrete*, and the terms in the second box pick out the *abstract* items in the world (if they pick out anything at all). Hence, according to van Inwagen, being "a platonist" (i.e., being someone who believes there are properties, and that properties are abstract, and hence non-spatiotemporal, acausal and imperceptible) "is [to be] someone who thinks that at least some of the linguistic items in the second class really are terms (really are predicates) and really have referents (really have non-empty extensions)" (ibid., 109).

Suppose we accept this way of subdividing whatever there is into 'the abstract' and 'the concrete'. Does this explain why van Inwagen refuses to countenance properties either in the sense of immanent universals or in the sense of tropes? I don't think so. For this way of distinguishing the abstract from the concrete can go 'wrong' (from the point of view of what van Inwagen thinks is true of objects and properties respectively) in many different ways. As we have seen, van Inwagen is well aware that, how we choose to categorize something will in part depend on our general philosophical background. This is why (Van Inwagen (2004), p. 109):

> …[m]ost philosophers would put '{Catherine the Great, {the Eiffel Tower}}' in with 'property' and 'the lion'; but the author of Parts of Classes might be inclined to think that this term was more at home with 'Catherine the Great' and 'the Eiffel Tower'.

And (ibid.):

> Amie Thomasson would say that our whole scheme of classification was in at least one respect objectionable, since 'War and Peace' isn't a clear candidate for membership in either class – for it denotes an object that is non-spatial and has instances (like many of the items in the second list), and is, nevertheless, a contingently existing artifact (like some of the items in the first).

And (ibid.):

> Nicholas Wolterstorff would say that our classification scheme was unobjectionable, and that 'War and Peace' clearly belonged right where we had put it, since it denoted something that was much more like a proposition than it was like a volume on a library shelf. He would add that the idea of a contingently existing, non-spatial object that had instances was incoherent.

What makes the difference between placing e.g., sets or 'War and Peace' in the abstract or in the concrete box, it seems, is our understanding of the nature of those items coming in. But, then, just as whether we come to the table with an understanding of sets as classes or not makes a difference to in which 'box' we think they belong, whether we understand properties as Platonic Forms or as tropes or immanent universals also matters to our categorizing them as either abstract or concrete. Rather than functioning as a comparatively speaking neutral starting-point for our categorizations, then, the putting into boxes imagined by van Inwagen seems to depend on preconceived ideas we have about the nature of the things we thus categorize.

Could it be that if most people think properties belong in the abstract box, this is because properties have something in common with the other items we tend to categorize as abstract? And could it be that the items we tend to categorize as abstract are typically non-spatiotemporal, acausal and imperceptible? Perhaps. But far from obviously. Van Inwagen places God, angels and elfs in the concrete box. But most seem to agree that neither of these are spatiotemporal entities. And he puts sentences – "as in 'the same offensive sentence … scrawled on every blackboard in the building" (Van Inwagen (2004), p. 108) – in the abstract box, although sentences, given that they can be offensive, seem to possess some sort of causal power. Even if properties are abstract, therefore, this does not obviously license our drawing any particular conclusions about their nature. And so it is not clear why it follows from their being abstract that they are non-spatiotemporal, acausal, and imperceptible. It doesn't seem that they *have to* be. For all we know, a van Inwagean Quinean *could* think of properties as spatiotemporally located, causally powerful and perceptible.

The point is not that Van Inwagen is wrong to categorize properties as abstract, or that he is wrong when thinking that this means that properties are, among other things, non-spatiotemporal, acausal and imperceptible. The point is only that from the point of view of his own project, it *makes sense* to say that properties are concrete, or that they are abstract, yet such that they exist in space and time, have causal powers, and can be perceived. But this means that the reason constitutionalism$_+$ doesn't make any sense – presuming it doesn't – *cannot* be that properties cannot be parts of or make up objects. And as we have seen earlier, it

cannot be because the ontological structure constitutionalism invokes must be inherently explanatory. But then it is unclear *why* constitutionalism doesn't make any sense. In fact, in view of the more scientifically respectable understanding of properties constitutionalism invokes, it seems we have van Inwagean reasons to think that it does.

7 End

According to van Inwagen, metaphysicians ought not to explain and they ought not to be constitutionalists. In this paper, I have argued that accepting the former claim does not mean having to accept the latter. In fact, constitutionalism$_+$ – the view that objects exist and are made up from properties – makes perfect sense, because ontological structure does not have to be inherently explanatory structure, and because properties do not have to be understood as non-spatiotemporal, acausal and imperceptible.

Acknowledgments: I'd like to thank Johan Brännmark, Peter van Inwagen and the audience at the 2017 *Quo Vadis, Metaphysics* conference, organized by Mirosław Szatkowski, for comments and criticisms of earlier drafts of this paper. This research was completed with the help of generous financial support from Riksbankens Jubileumsfond and the University of Gothenburg.

Bibliography

Armstrong, D. M. (1978), *Universals and Scientific Realism* I-II, Cambridge UP.
Armstrong, D. M. (1997), *A World of States of Affairs*, Cambridge UP.
Armstrong, D. M. (1997/1980), "Against 'Ostrich' Nominalism: A Reply to Michael Devitt", in *Properties*, edited by H. Mellor and A. Oliver, Oxford UP, 101–111.
Audi, P. (2012), "Grounding: Toward a Theory of the In-Virtue-Of Relation", in *The Journal of Philosophy*: 109, 685–711.
Bergmann, G. (1967), *Realism*: *A Critique of Brentano and Meinong*, University of Wisconsin Press.
Dasgupta, S. (2017), "Constitutive Explanation", in *Philosophical Issues*: 27(1), 74–97.
deRosset, L. (2013), "Grounding Explanations", in *Philosophers' Imprint*: 13(7), 1–26.
Fine, K. (2012), "Guide to Ground", in *Metaphysical Grounding*: *Understanding the Structure of Reality*, edited by F. Correia and B. Schnieder, Cambridge UP, 37–80.
Lewis, D. (1983), "New Work for a Theory of Universals", in *Australasian Journal of Philosophy*: 61(4), 343–377.

Lewis, D. (1996), "Elusive Knowledge", in *Australasian Journal of Philosophy*: 74(4), 549–567.
Loux, M. J. (2017), "Theories of Character", in *Being, Freedom and Method: Themes from the Philosophy of Peter van Inwagen*, edited by J. A. Keller, Oxford UP, 11–31.
Nolan, D. (2005), *David Lewis*, Acumen Publishing.
Paul, L. (2002), "Logical Parts", in *Noûs*: 36(4), 578–596.
Paul, L. (2017), "A One Category Ontology", in *Being, Freedom and Method: Themes from the Philosophy of Peter van Inwagen*, edited by J. A. Keller, Oxford UP, 32–61.
Quine, W. V. O. (1948), "On What There Is", in *The Review of Metaphysics*: 2(5), 21–38.
Raven, M. J. (2015), "Ground", in *Philosophy Compass*: 10(5), 322–333.
Van Inwagen, P. (2004), "A Theory of Properties", in *Oxford Studies in Metaphysics*, 1, edited by D. W. Zimmerman, Clarendon Press, 107–138.
Van Inwagen, P. (2011), "Relational vs. Constituent Ontologies", in *Philosophical Perspectives*: 25(1), 389–405.
Van Inwagen, P. (2014a), "Dispensing with Ontological Levels: An Illustration", in *Disputatio*: 6, 25–43.
Van Inwagen, P. (2014b), "What is an Ontological Category", in *Existence*: *Essays in Ontology*, Cambridge UP, 183–201.
Van Inwagen, P. (2017), "Reply to Michael J. Loux", in *Being, Freedom and Method: Themes from the Philosophy of Peter van Inwagen*, edited by J. A. Keller, Oxford UP, 344–348.
Wolterstorff, N. (1970), "Bergmann's Constituent Ontology", in *Noûs*: 4(2), 109–134.
Wolterstorff, N. (1991), "Divine Simplicity", in *Philosophical Perspectives*: 5, 531–552.

Carl J. Posy

Inside the Metaphysical Workshop

Abstract: The paper begins by noting that a certain ontology (van Inwagen's favored ontology) seems naturally to invite a formal semantic grounding (that is support from contemporary formal model theory). Using variations on classic and recent meta-semantic arguments, the paper goes on to show that such a grounding is not available. It suggests that the reason for this failure is that formal semantics is the wrong tool for the task of grounding particular ontologies. Building on that negative argument, the paper concludes by suggesting that formal model theory is well suited for a different metaphysical task.

1 Introductory remarks

This essay's point is philosophical, really meta-philosophical: I am going to argue against one tempting use of formal semantics in metaphysics, and at the end in favor of a different use. But I'm going to make these points via a technical exercise: I'm going to apply a variation on Ori Simchen's notion of scrambled truth[1] – itself a variation on Putnam's model theoretic argument – to a meta-metaphysical question about the use of formal semantics. My argument will go weird at one point. But I expect that this is acceptable. I'm learning that contemporary metaphysics can sometimes be outré.

I say "I'm learning"; for, I'm a newcomer to contemporary metaphysics. The present contribution to this volume honoring Peter van Inwagen marks my tentative first steps in the big-time after apprenticeships in logic and the history of philosophy.

0 I would like to thank Ori Simchen for very helpful correspondence on the topics of this paper and Bar Luzon for invaluable discussions at the initial stages and for her aid in tracking down a number of bibliographic sources. Discussions with Prof. Van Inwagen at the Warsaw conference (*Quo Vadis Metaphysics*) were very important. I also had helpful feedback from the audiences at Victoria University of Wellington, Canterbury University in Christchurch, and the University of Waikato in Hamilton, New Zealand. Miroslaw Szatkowski and Rea Golan made some crucial suggestions for improving the paper. I owe a special debt of gratitude to Professor Szatkowski for his encouragement and patience, without either of which this paper would not exist. Grant 1491-14 from the Israel Science Foundation funded the research for and writing of this paper.

1 Set out in Simchen (2017a,b).

https://doi.org/10.1515/9783110664812-008

That's an oblique allusion to my title. The idea of a 'metaphysical workshop' connotes a guild of masters and apprentices. Here's the explicit point: van Inwagen speaks of the 'ontology room'[2] (a refinement of David Lewis' "philosophy room"); and I changed 'room' to 'workshop' to emphasize that we have certain tools and products. I want to talk about a tempting but inappropriate use of a certain tool to produce a certain product, and to speculate upon its proper use. I believe that in metaphysics – in philosophy in general – our tools should suit our goals and that we would do well to avoid the temptation to apply those tools beyond their intended scopes.

"Formal semantics" – the abstract account of truth, that rests on the logical form of truth bearers (propositions or sentences) – is, as I said, the tool I will discuss. The tempting product is what van Inwagen calls an "ontology"; or more precisely, the product is to be a justified ontology.

Van Inwagen gives detailed criteria for being an ontology, but for our purposes initially the important points are that an ontology is an exhaustive taxonomy of non-empty, non-overlapping, and unified categories of objects. *Exhaustive*: Everything that *is* falls into one or another ontological category. *Non-overlapping*: There are no multiple memberships. *Unified*: A category is never a random collection. There is always something in virtue of which denizens of a category live in that category and not elsewhere.[3] *Justified*: Metaphysicians don't just think up candidate ontologies and then leave. They try to support their proffered ontologies on reasonable grounds. Grounds that show, in particular, that the categories of those ontologies do apply to worldly things, are truly unified and truly exhaustive and do not overlap.

For the most part, I'm going to concentrate on a particular candidate ontology. It is the one favored by van Inwagen himself, and it consists of two basic categories: individual substances, and relations (including propositions and monadic properties).[4] These latter are what van Inwagen sometimes calls "assertibles", or

2 "[I]n philosophy's house there are many rooms, and one of them, more austere in design and more sparsely furnished than perhaps any of the others, is the ontology room." (Van Inwagen (2014b)).

3 It is an assumption of ontology that there are natural classes whose membership comprises a really significant proportion of the things that there are. … [T]he existence of natural classes follows from the existence of real divisions among things. Nevertheless, the concept of 'natural class' cannot be defined solely in terms the concept 'real division'. We must also appeal to the concept of 'sufficient internal unity' if we are to provide a full explanation of 'natural class'." (Van Inwagen (2012)).

4 According to [the ontology I myself favor] there are two primary categories, substance and relation. … The category 'relation' subsumes propositions (0-adic relations) and attributes (monadic

"unsaturated objects".[5] This is an ontology: van Inwagen claims that everything that *is* belongs to one or the other of these classes, that the classes don't overlap, and that each of them is unified. I'll call this *VIFO*, for Van Inwagen's Favored Ontology. I chose it not only to get a pronounceable acronym, but because it is the most natural and most tempting candidate for a semantic grounding (a justification or a derivation). For, *VIFO* corresponds closely to the basics of the modern conception of logical form and semantics – the conception codified in contemporary model theory.

There is a longstanding tradition of using logical form (in particular the form of basic predications) to ground ontology.[6] For our purposes I will call attention to Leibniz. For, Leibniz – at least according to Bertrand Russell – justified his ontology of monads by an appeal to his own formal semantics, his containment theory of truth.[7] Russell lambasted that theory, mainly because it rested on outdated Aristotelian logical grammar: basic predication was a monadic general term applied to another such term, and quantifiers were part of the terms.[8] Russell in fact helped revamp that grammar into the modern conception of logical form and the attendant formal semantics.

Specifically: On the side of logical grammar, we now believe that predication is the application of a monadic or relational general term to a singular term (or terms), and we hold that quantifiers are independent operators. This is what we use to build our formal systems. On the semantic side, we define structures to model those systems and to exhibit their logical truths and their consequences. Tarski gave this enterprise its most precise form. He introduced the precise notion of a metalanguage in which we discuss the references of our terms and the truth of our sentences; he codified the recursive definition of reference itself, and he introduced the satisfaction relation in order to provide a workable semantics for quantifiers. This is the heart of model theory.

---

relations). The category 'substance' goes by two other names, "concrete thing" and "individual (thing)". Similarly, the category 'relation' is also called 'abstract thing' and 'universal'." (Van Inwagen (2012)).

5 "I propose, therefore, that properties be identified with unsaturated assertibles, with things that can be said of things. It seems unproblematical that unsaturated assertibles can successfully play the property role. And I would ask this: what is the property of whiteness but something we, in speaking of things, occasionally predicate of some of them? And what is predicating something of something but *saying* the former *of* the latter." (Van Inwagen (2004)).

6 Cocchiarella (2001) gives a nice summary of historical cases.

7 See Russell (1937).

8 See Russell (1938), in particular e.g. §215 and §449.

I tell you these things, because once our modern theory is in place, the argument from model theory to *VIFO* seems a slam-dunk: truth mirrors the way the world is; and the truth of statements rests ultimately on the satisfaction of basic predications – the application of a predicate term to a singular term or singular terms – the base of logical form.[9] The referents of that singular terms will be individuals and the referent of the general term will be a relation. Voila! Simply read off from the form of your judgment the categories of the things that make that judgment true and note that this logical form is basic and general.

Now there certainly are authors who embrace this affinity between model theory and ontology: Frege's ontology of objects and functions fits the bill. Russell himself favored an ontology close to *VIFO* and at least at some points seemed to connect his own favored ontology to that updated formal semantics[10] P. F. Strawson also played on this parallelism between logical form, truth and ontology.[11] More recently, Nino Cocchiarella generalized this approach.[12] Even more recently than that Theodore Sider has argued for the ontological significance of logical structure and quantification.[13]

Van Inwagen seems to take his place in this company. Two things he says give the very clear impression that he indeed grounds *VIFO* on model theory. First, following Quine, he says that we must codify what we know of the world in a first order language. Secondly, he says that we derive the categories by classifying the "satisfiers" of our satisfied sentences.[14] Indeed, at one point he says that we do

9 This is a point that Strawson (1974) argues for powerfully. He also refers to Quine and Chomsky as fellow travelers on this. See also Cocchiarella (1987), ch. 2.

10 "There are particulars and qualities and relations of various orders, a whole hierarchy of different sorts of simples, but all of them, if we were right, have in their various ways some kind of reality that does not belong to anything else." (Russell (1918), p. 11).

"… The purpose of the foregoing discussion of an ideal logical language (which would of course be wholly useless for daily life) is twofold: first, to prevent inferences from the nature of language to the nature of the world, which are fallacious because they depend upon the logical defects of language; secondly, to suggest, by inquiring what logic requires of a language which is to avoid contradiction, what sort of a structure we may reasonably suppose the world to have." (Russell op. cit., pp. 144–145).

11 This is the core thesis of his *Subject and Predicate in Logic and Philosophy*. Though he aims to reverse the direction of grounding (he argues from ontological categories to logical form), nevertheless his argument rests on the same general procedure. And, in fact occasionally he presupposes (malgré lui) the direction from logical form to ontological category.

12 See Cocchiarella (1996, 2001, 2007).

13 See Sider (2011), especially Chapter 6.

14 "Look at all the things that you, the ontologian, believe 'outside' ontology – the beliefs that, as it were, you bring to ontology. Subject them to quantificational analysis a' la Quine. This will provide you with a large class of one-place open sentences that you believe are satisfied. Try to

this ontological derivation in a language he calls "Tarskian".[15] It seems that he is explicitly providing a framework for a semantic grounding of *VIFO*.

Now, in fact – "Tarskian" and "satisfiers" notwithstanding – van Inwagen does not rest his defense of *VIFO* on model theory.[16] And I believe that this is a good thing for him. For, when you fill in the needed details, you'll see that the grounding fails. Simchen style 'scrambling' will deliver the coup de grace. Ultimately, I'll say, the attempted grounding fails because formal semantics is an inappropriate tool for this task. I believe the reasons for that 'inappropriateness' are in fact deep, or at least interesting.

So in the first part of this paper, I'm going to show you what we would have to do if we wanted to rest *VIFO* on formal semantics, on model theory. I will indeed use some of the things van Inwagen says as if they were instructions for effecting such a grounding. In the course of doing this I will show you that if there is any chance to make the purported semantic grounding of *VIFO* work, then we must view the formal semantics in a particular meta-semantic light. (We will need to adopt the doctrine of 'reference magnetism'.) But then, using variations on Simchen's technique, I'll show you that, even so, this grounding won't work. We can't derive *VIFO* from model theory or use model theory to justify *VIFO*. As I said, I will argue that this is so because model theory is a tool for a different project altogether; and, in our workshop, tools must suit their uses. I will use this argument to give a gloss on van Inwagen's semantic-seeming remarks. But I'll conclude by asking whether there is a proper place for model theory in the metaphysician's tool kit, and I'll give a speculative yes answer. I will say in fact that the very same aspect of model theory that disqualifies it as a tool to ground an ontology serves instead as a means to express an entirely different metaphysical product, a product I shall call 'modal modesty'. I'll conclude with a very brief historical allusion designed to show that this 'modesty' too has a place inside the metaphysical workshop.

---

give a coherent account of the 'satisfiers' of those sentences, a project that will, in some cases, involve fitting them into a system of ontological categories. See whether the resulting system of categories satisfies you intellectually. Subject it to all the dialectical pressures you can muster – and attend to the dialectical pressures those who disagree with you bring against it." (Van Inwagen (2011)).

15 "[P]articipants in discussions in the ontology room … converse in a language I will call Tarskian." (Van Inwagen (2014b)).

16 He said so explicitly during the 2017 Warsaw conference in his honor, *Quo Vadis Metaphysics*.

2 A purported model theoretic justification of *VIFO*

So let's see what it would take to justify *VIFO* as the right ontology using our basic observations about the world and our formal (model theoretic) notions of grammar, reference, and truth.[17]

2.1 Assumptions and principles

To get started I'll indicate the working assumptions that we can derive from van Inwagen's remarks and then refine them in order to streamline the purported grounding of *VIFO*. Then I'll highlight the points we need from formal model theory.

Here is the pair of working assumptions gleaned from Van Inwagen:

- We assume that we have described the world in a simple first order language $\mathcal{L}$. $\mathcal{L}$ has predicate terms and names for objects.[18] It makes sentences out of these using the standard logical operators. Thus there will be a set, Σ, of sentences in $\mathcal{L}$ that we take as truly describing the world. These are the *facts on the ground*, the *data*. The objects in the world will be the satisfiers of the satisfied open sentences in Σ.
- We insist that we can ultimately describe whatever there is in the world. There are, I shall assume, no ghostly beings that in principle elude our best scientific efforts to depict reality. So a language like $\mathcal{L}$ suffices to describe what there is. This isn't physicalism, but it is a *deep empiricism*. It says we can in principle name and know whatever things there are, even if we haven't in fact named them all or discovered all there is to know about them. Though metaphysical sounding, this empiricism simply says that $\mathcal{L}$ can describe everything there is.

17 Van Inwagen's full version of *VIFO* is distinguished from other ontologies in the literature in virtue of being what he calls "relational": "There is no possible sense of 'constituent' in which an abstract object [i.e., a property – CP] can be a constituent of a substance/concrete particular/individual thing." (Van Inwagen (2012)). It would be an interesting exercise to see if relationality is also a candidate for semantic or model theoretic grounding. From the present point of view, this is of course ultimately moot. I am arguing against a semantic grounding for even the most tempting aspects of an ontology. See, however, note 21 below.

18 Some authors – Quine, most prominently – argue for eliminating individual names in favor of predicates. I will not make this move. For, among other reasons, individual designators remain crucial when we implement the clauses for satisfaction in a formal semantics.

Here are four refinements of those assumptions designed to streamline our purported grounding of *VIFO*:

- *I will assume that for any property that applies to anything in the world, $\mathcal{L}$ contains at least one name of something to which that property applies*. There are endlessly many worldly individual things: individual crayfish, barnacles, belt buckles and buildings to name a few. Well perhaps not *endlessly* many if we restrict it to concrete things, but certainly a large finite number of them, even if we don't manage to name them all. My additional assumption simply takes our empiricism seriously. It allows us to bypass modifications that arise as our language evolves.[19] But it also has a technical advantage that I'll mention below.
- *I will assume that $\mathcal{L}$ has no purely semantic terms*. Van Inwagen will insist that $\mathcal{L}$ should have no purely metaphysical terms. (Our task to derive metaphysics; we don't assume it.) Since we are working on the hypothesis of the metaphysical significance of semantics, it is best also to assume that $\mathcal{L}$ has no purely semantic terms either.[20] I won't say much about what constitutes metaphysical or semantic discourse. Suffice to say that classifying something as an ontological category is a metaphysical statement, as is speaking of its ontological characteristics. Similarly, properties like being a truth bearer or being true and relations like referring or naming are semantical.
- Worldly things have many relations; but I will concentrate only on monadic properties. Thus, *I will assume that $\mathcal{L}$ is a monadic first order language*.[21]
- *I will further restrict attention to first order properties*. That is, I will deflect consideration of properties of properties.

I make these additional refinements in order to concentrate on the essential points of my argument. It is subtle enough at places, and I want to avoid distracting complications. I am persuaded that nothing crucial to the argument rests on them. I will however indicate briefly below the modifications that we would need to make in order to accommodate polyadic relations and higher order properties.

19 This is not a constructivism. It does not place restrictions on the grounds for asserting an existential claim.

20 In general, the difficulties that I am raising for an alleged model theoretic grounding of *VIFO* are not those of the semantic paradoxes. Nonetheless, we want to be sure to steer clear of such pitfalls. Thus, in particular, given the assumption about namability the proviso that $\mathcal{L}$ be semantics-free will serve to avoid worry about such paradoxical properties as "not being named in $\mathcal{L}$".

21 To remarks: First, keeping *VIFO* monadic sidesteps issues of identity – identity of substances and identity of properties. Were we to have identity statements, they would be subject to the mischief that I will describe below. Secondly, were we to allow $\mathcal{L}$ to be a polyadic language, some of these would be empirical claims about how some things are composed of others. ("The nib is part of the pen".) It will also include meaning relations among predicates. ("Xenophobia means excessive fear of strangers".) Claims of these two kinds would be important for an argument establishing what van Inwagen calls the 'relationality' of *VIFO*.

I also add a practical *ad hoc* assumption:

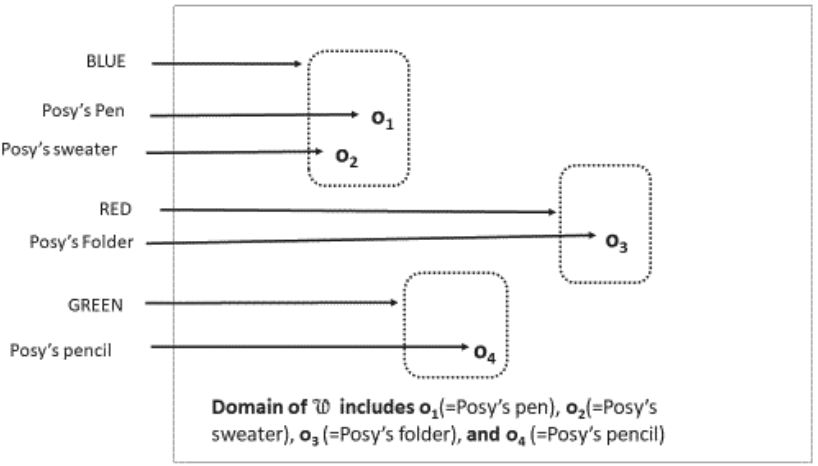
- It happens that I own a blue pen, a blue sweater, a red document folder, and a green mechanical pencil. So the expressions "Posy's pen", "Posy's sweater", "Posy's folder", "Posy's pencil" will be among the names of $\mathcal{L}$, and "BLUE", "RED" and "GREEN" will be among the predicate terms.

Finally, here are the technical points we need:

- We can assume Σ is consistent and non-empty. So, Σ has a model, $\mathcal{W}$. That is a basic fact of model theory. The Appendix has a version of the technical details of basic model theory. For now I will emphasize only two points:
  - There is domain set, $D_{\mathcal{W}}$, of objects and an interpretation $\mathcal{I}^{\mathcal{W}}$ of the language $\mathcal{L}$ that assigns objects in $D_{\mathcal{W}}$ to the names in $\mathcal{L}$, and assigns subsets of $D_{\mathcal{W}}$ to the general terms of $\mathcal{L}$.
  - This interpretation makes all of the elementary predications in Σ (all the sentences of the form $Q_i(a_n)$) come out true. (This is principle T1 in the Appendix.) Following this, the recursive definition of truth will make all of the compound sentences of Σ come out true in $\mathcal{W}$.
- Since Σ is about the world, we of course assume the objects of $\mathcal{W}$ are worldly objects. So following my ad hoc assumption, my pen, sweater, folder and pencil are in the domain, and they have their worldly properties. I will speak of them in our metalanguage as o_1, o_2, o_3, and o_4 respectively.

Figure 1 depicts the main parts of all this:

Figure 1: **Reference of Terms in the Model $\mathfrak{W}$**



The boxes depicted are subsets of the extensions of "Blue", "Red" and "Green". (Recall there will be lots more names than these in $\mathcal{L}$.) Now apply the model theoretic rules of truth, and you will get the truth in $\mathcal{W}$ of sentences such as: *Posy's pen is blue, so is his sweater; but his pencil is green, and his folder is red*.

2.2 The argument

We aim to establish, then – using only our working assumptions, together with Σ, $\mathcal{W}$ and the principles of model theory – that there really are such things as individuals and properties, and that these two classes of objects do in fact compose an ontology – at the very least that they are non-overlapping, exhaustive and unified.

2.2.1 The categories are non-empty

First, to show that those classes – individuals and properties – are non-empty, let me remark again that the heart of the matter is elementary predication, in our monadic case the application of a predicate to an individual. Predication is the fulcrum of modern semantics, the place where the rubber hits the road, where language and world first come together to produce truth.[22] The truth clause for predication is quite explicit: $Q(a)$ is true just in case the interpretation of a, $(\mathcal{I}(a))$, is an element of the interpretation of Q, $(\mathcal{I}(Q))$. (This is principle TI in the Appendix.)

As I said, those rounded boxes in Figure 1 represent sets; but let me remark that nothing rests on set theory here. Model theory does not say that properties are sets or that property application is identical to set membership. In fact, it says nothing at all about what a worldly property is, nor about how application works. The formal model theoretic truth conditions simply present the *extensional* outcome of applying a property to an individual. This is a limited representation of predication and its consequences. It is, however, precisely the level of representation we need for a logical grounding of *VIFO*.[23]

Having said this, we have completed our first task: There are individuals in the domain of $\mathcal{W}$ (predica*tees*); and whatever it is that has the set $\mathcal{I}(Q)$ as its extension,

22 Indeed, this relation remains fundamental even if we do eliminate names in favor of predicates, and even – if like Sider (2011) – we were to attribute categorial ontological significance to quantified expressions. For, the predication relation shows up in the satisfaction conditions for quantified statements.

23 That caveat jibes well with van Inwagen's view of these things. For, Van Inwagen has no truck with properties as sets, and he is explicitly agnostic about the actual nature of application.

that predica*tor* is a property. Thus the two classes – individuals, and properties – are, as the ontology claims, non-empty.

2.2.2 The predicative profile, $\mathcal{W}^{\#}$, of $\mathcal{W}$

We still need to show that these two classes are exhaustive, unified and non-overlapping; that is that they can constitute an ontology. To do this I introduce what I call the "*predicative profile*", $\mathcal{W}^{\#}$ of $\mathcal{W}$. $\mathcal{W}^{\#}$ lets model theory cast its clinical spotlight on predication: It "x-ray's" the model, $\mathcal{W}$, and presents that model's predicative skeleton in relief.

Technically, $\mathcal{W}^{\#}$ is itself a model. This, effectively, is the place in our argument where model theory become explicitly metaphysical. Here are the details of $\mathcal{W}^{\#}$:

The language, $\mathcal{L}^{\#}$ that this model interprets contains names for each of the symbols in the vocabulary of $\mathcal{L}$ together with the binary general term "A" and the two unary general terms "S" and "P".

The domain, $D_{\mathcal{W}^{\#}}$, will contain three kinds of objects:

- First, all the elements in the domain of $\mathcal{W}$ that are designated by names of $\mathcal{L}$;
- Second, all the properties that truly apply to at least one of those element in $\mathcal{W}$. As Van Inwagen insists, from a metaphysical point of view, properties are objects too. Thus in particular, the properties 'being blue', 'being red' and 'being green' are objects in the domain of $\mathcal{W}^{\#}$. Let's call them o_5, o_6 and o_7 respectively
- Third, the object, $\perp$.
  - This is the "empty property"; that is, the property that applies to nothing in $D_{\mathcal{W}}$. For our purposes it is sufficient to conflate all the properties with empty extensions in $\mathcal{W}$.

"There is only one objection to Lewis' theory of properties [i.e., properties are sets – CP]; it isn't true." (Van Inwagen (2004))

"The Lewis-Heidegger problem may be framed as a question: 'How does a certain concrete object (a green ball, for example) reach out and take hold of a certain propostion (the proposition that there is at least one green ball, for example), an abstract object, and make it true? The question, 'How does a concrete object (like a green ball) reach out and take hold of a propoerty (like the color green), an abstract object, and make it had or instantiated? is at least a very similar question. ... In my opinion, these questions have no answer ... " (Van Inwagen (2011))

- Here's where our extended empiricism kicks in, the assumption that if a property applies to something, then we can name at least one thing to which the property applies. The properties consolidated into $\perp$, are truly empty. None of them is an instantiated property that happens accidentally to apply to nothing that is named.

The individuals and the properties that we have put into $D_{\mathcal{W}^{\#}}$ swim around together in there on an equal footing. The domain is not explicitly "typed."

The interpretation function, $\mathcal{I}^{\mathcal{W}^{\#}}$, works as follows:

- The interpretation of an $\mathcal{L}^{\#}$ name of the form "'a_i'" is the entity to which 'a_i' refers in $\mathcal{W}$.
- The interpretation of an $\mathcal{L}^{\#}$ name of the form "'Q_i'" is the property that determines the extension, $\mathcal{I}^{\mathcal{W}}(Q_i)$, in $\mathcal{W}$. (Keep in mind that we remain agnostic about how that determination is made.)
- $\mathcal{I}^{\mathcal{W}^{\#}}(A) = \{< \mathcal{I}^{\mathcal{W}^{\#}}(Q_i), \mathcal{I}^{\mathcal{W}^{\#}}(t) > \mid \mathcal{W} \models_S Q_i(t)\}$.
  - Here's place where my assumption that $\mathcal{L}$ is monadic makes a difference. For, if $\mathcal{L}$ were to be a polyadic language, then A would have a variable arity. We might have an object o that finds itself sometimes in the lead place of an n-tuple and sometimes of an m-tuple where $m \neq n$.[24]
- $\mathcal{I}^{\mathcal{W}^{\#}}(P) = \{\mathcal{I}^{\mathcal{W}^{\#}}(Q_i) \mid \text{for some } o, < \mathcal{I}^{\mathcal{W}^{\#}}(Q_i), o >\in \mathcal{I}^{\mathcal{W}^{\#}}(A)\} \cup \{\perp\}$.
- $\mathcal{I}^{\mathcal{W}^{\#}}(S) = \{o \mid \text{for some } i, < \mathcal{W}^{\#}(Q_i), o >\in \mathcal{I}^{\mathcal{W}^{\#}}(A)\}$.
  - Here's where the restriction to first order properties kicks in. Were we to admit higher order properties, then the extension of S would have to be a bit more complicated: $\mathcal{I}^{\mathcal{W}^{\#}}(S) = \{o \mid \text{for some } i, < \mathcal{W}^{\#}(Q_i), o >\in \mathcal{I}^{\mathcal{W}^{\#}}(A)$, and for no object o' is it the case that $< o, o' >\in \mathcal{I}^{\mathcal{W}^{\#}}(A)\}$.

Informally, $A(x,y)$ says that x is a predicate term and y is a name in $\mathcal{W}$ and that the designated predicate applies to that named object. It says, for instance, that BLUE applies to my pen.[25]

---

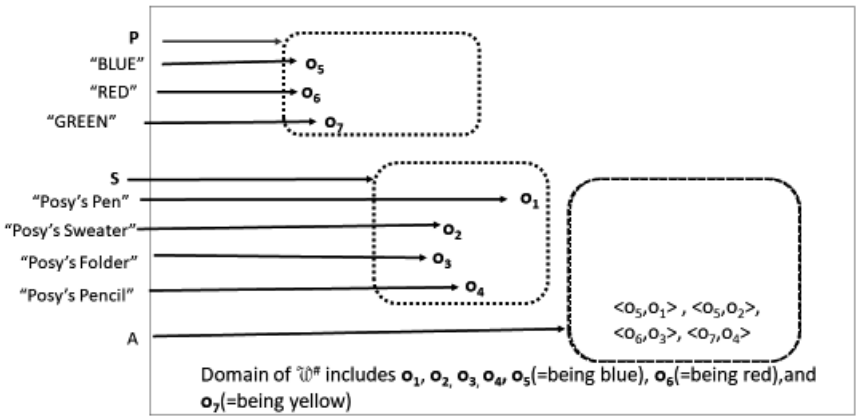
24 For both of these worries I refer you to Richard Grandy's "anadic logic". (See Grandy (1976) and Grandy (1977), chapter XIV). Grandy gives a logic and semantics geared to deal with the fact that our ordinary properties are often not arity-specific. Consider, "live together in an apartment". The adaptation of our treatment to such a logic, is a mechanical exercise. It is natural, and in fact corresponds to a number of van Inwagen's remarks.

25 A word about the predicate $A(x,y)$, the predicate expressing the relation of predication itself. Since we are viewing $A(x,y)$ as a proposition, we are led to consider the profile $\mathcal{W}^{\#\#}$ in which this predication itself is highlighted. This clearly leads to an infinite progression. The classic "problem of predication", or the problem of the "unity of the proposition" as one occasionally puts it, is the question of what makes predications work in the first place. The difficulty is that such an infinite chain cannot serve to provide the metaphysical or semantic "glue" that converts a mere juxtaposition of concepts or predicates into a unified, truth-value bearing, proposition. If our goal were to explain the nature or possibility of predication and truth, then indeed this would be

The extensions of S and P, in $\mathcal{W}^{\#}$, are the set of second members and the set of first members respectively in the extension of A. S and P designate substances and properties respectively. (If we were to adopt the second version of S that I mentioned above, then S's extension would be the set of second members of ordered pairs in $\mathcal{I}^{\mathcal{W}^{\#}}(A)$ such they never appear as a first member. This is close to the Aristotelean notion of a substance.)

Here's a picture to show how all this works.

Figure 2: **Reference of Terms in the Model $\mathfrak{W}^{\#}$**



Notice that the predicate, A, reflects in $\mathcal{W}^{\#}$ what's going on in $\mathcal{W}$ in the following way.

PRINCIPLE #: $\mathcal{W}^{\#} \models A(\text{“}Q_k\text{”},\text{“}t\text{”})$ just in case $\mathcal{W}^{\#} \models Q_k(t)$.[26]

2.3 The argument continued

Now it is straightforward to show what we want. We need to show: (i) that each of S and P is a truly unified collection. (ii) that they are exhaustive – there are no

a problem. But that is not our goal — neither in model theory nor in the business of grounding ontologies. Thus for our purposes, there is no problem at all with generating such an infinite chain.

26 Notice that since all the elements of $D_{\mathcal{W}}$ are named in $\mathcal{L}^{\#}$, we can dispense with assignments and define quantified truth directly.

unnamable gremlins lurking in the domain that that might couple in unnamable ways –, and (iii) that they don't overlap.[27]

Well, S and P are truly unified: being a first member in a predication is a pretty unifying trait; so too for being a second member. Our assumed empiricism for $\mathcal{W}$ banishes any worrisome gremlins in $\mathcal{W}^{\#}$. Finally, our grammar gives us the needed duality. In a regimented language such as $\mathcal{L}$ things know their place. The collection of singular and general terms is fixed and distinguishable, and thus first and second places in A are as well. There is no overlap.

So far so good, then. This is how the formal semantic justification of *VIFO* seems to go. We use $\mathcal{W}$ to generate the categories Substance and Property, we use empiricism to show that they are exhaustive. Then we use $\mathcal{W}^{\#}$ to show them distinct and unified. So it would seem that from the facts on the ground and formal semantics – buttressed by empiricism – we can indeed ground *VIFO*.

3 Mischievous permutations

3.1 The model theoretic argument

But putting things this starkly highlights where trouble lurks. It is what we might call the loose grip of the reference function $\mathcal{I}$. That is the point of Putnam's model theoretic argument. Let me remind you of how that argument would go in our special case: We have Σ, our set of sentences that are true in $\mathcal{W}$ (that is true of the world). Putnam shows how to build an entirely different model, $\mathcal{W}^{\star}$ of Σ, made from the very same objects as $D_{\mathcal{W}}$, whose interpretation function assigns different extensions to the predicates. He gets this by first permuting the domain so that $\mathcal{L}$'s terms refer to different things.[28] Then he gerrymanders the interpretations of the general terms so that same atomic sentences come out true in $\mathcal{W}^{\star}$ as did in $\mathcal{W}$.

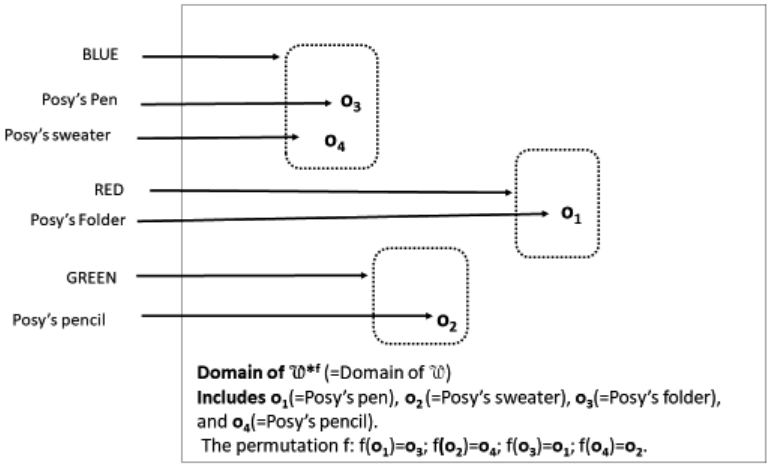
Here's an example: Suppose after the permutation the term "Posy's pen" ends up refering in $\mathcal{W}^{\star}$ to my red folder (object o_3). Thus to keep the sentence "Posy's pen is blue" as true, we need to put that folder into the extension of 'Blue'. The same can go for other predicates. To fill out the example, suppose the permutation f is defined as follows: $f(o_1) = o_3$; $f(o_2) = o_4$; $f(o_3) = o_1$; $f(o_4) = o_2$. Let's call the resulting model $\mathcal{W}^{\star f}$. Here's a picture of how we would gerrymander the predicate-

27 Recall, we are dealing only with first order properties.

28 Think of the elements of $D_{\mathcal{W}}$ as ordered linearly; so a given name 'a' denotes an item that stands in a particular place in the ordering. The effect of the permutation is that 'a' now denotes whatever element of the domain now stands in that particular place in the ordering.

extensions in order to preserve the truth of the Σ sentences about those worldly objects.

Figure 3: **Reference of Terms in the Model $\mathfrak{W}^{*f}$**



This is Putnam in his anti-realist middle period. He is arguing against what he called "metaphysical realism", the view according to which reality is independent of our humanly constructed theories. In particular, the realist will say that even if Σ were to be an "epistemologically perfect" scientific description of the world, nevertheless it might be false of the world. Putnam's point is that it is always possible to build a model based on the objects of the world in order to verify a given theory.[29]

A defense of anti-realism, however, is not our current task; it is not the product that our metaphysical workshop has promised to provide. Identifying and justifying ontological categories is the promised product. In this task, however, Put-

29 "[L]et T_1 be an ideal theory, by our lights. Lifting restrictions on our actual all-too-finite powers, we can imagine T_1 to have every property *except objective truth* – which is left open – that we like. … to meet whatever 'operational constraints' there are … . The [Metaphysical Realist's – CP] supposition under consideration is that T_1 might still be all this and *still be* (in reality) *false*. … Pick a model $\mathcal{M}$ of the same cardinality as THE WORLD, map the individuals of M one-to-one into the pieces of the THE WORLD. The result is a satisfaction relation SAT – a 'correspondence' between the terms of [our language] $\mathcal{L}$ and sets of pieces of the THE WORLD – such that the theory T_1 comes out *true* – true of THE WORLD – provided we just interpret 'true' as TRUE(SAT). So what becomes of the claim that even the *ideal* theory T_1 might *really* be false?" (Putnam (1977a)). See also Putnam (1977b).

nam's 2-step jig is really quite mischievous. For, it leads to Rube Goldberg combinations that whammy our predicative profile, $\mathcal{W}^{\#}$, and our attempt to ground *VIFO*. For instance, we can build a model $\mathcal{W}^{\#*g}$ with the following permutation, g, of the domain, $D_{\mathcal{W}^{\#}}$: $g(o_1) = o_5$; $g(o_2) = o_6$; $g(o_3) = o_4$; $g(o_4) = o_3$; $g(o_5) = o_1$; $g(o_6) = o_2$; and $g(o_7) = o_7$. In words:

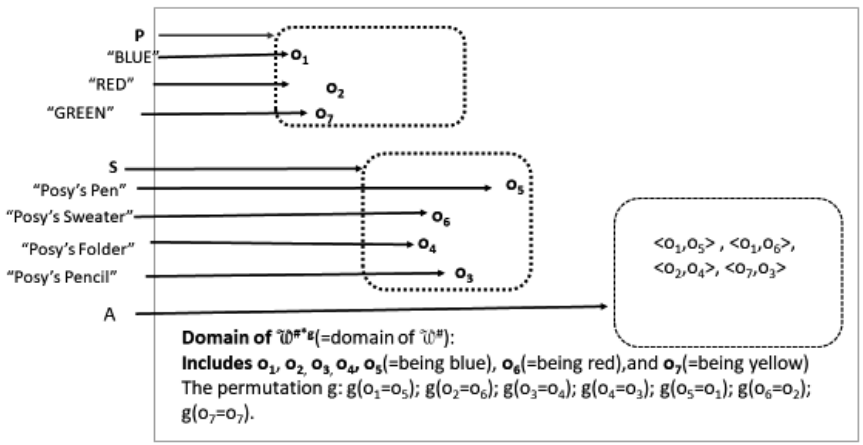
- The expression "Posy's pen" refers to the property of being blue.
- The expression "Posy's sweater" refers to the property of being red.
- The expression "Posy's folder" refers to my pencil.
- The expression "Posy's pencil" refers to my folder.
- The expression "Blue" refers to my pen.
- The expression "Red" refers to my sweater
- The expression "Green" refers to the property of being green

Given this permutation, we can still build a model that makes true all the predications of S and P that were true in $\mathcal{W}$.

Thus in order to preserve the truth of the sentence A("Blue","Posy's pen") we will have to predicate the pen of blue. Perhaps worse yet, in order to preserve the truth of the sentence A("Red", "Posy's folder"), we must predicate my pencil of my sweater. (This is a weird spot in the argument, as I promised.)

Once again, here's a chart that lays all this out graphically:

Figure 4: **Reference of Terms in the Model $\mathfrak{W}^{\#*g}$**



To be sure, we might concoct a story, in a Quine-like fashion, in which the object blue is an interrupted predicatee that pops up now and again when we predicate of it things like my sweater or my pen. And perhaps in some Whorfian way

you might indeed contemplate a culture in which my pen and my pencil are predicat*ors* while Blue and Red are predicat*ees*.[30] This, however, is scarce comfort for *VIFO* fans. For, the fact is that this bit of mischief really scotches any claim to unity that S and P might have. We simply threw them together by a random permutation of the domain. There are many, many such permutations. Think now of the endless or at least very large domain $D_{\mathcal{W}}$. Some – most, I daresay – of the random couplings that we can produce are so outlandish as to be absurd. A crayfish predicated of the property of being a pencil, a barnacle of being cow, my belt buckle of the Taj Mahal. They are, as one says, so strange as not to be even false. Odd (indeed monstrous) "ontologies" like these are really no more than the interaction of such a random permutation with the desire to maintain the truths of $\mathcal{W}^{\#}$.

3.2 Reference magnetism

The most prominent and most natural way that has been proposed to avoid these worries in general is David Lewis' response to Putnam: Lewis, rightly, points out that the Putnam gerrymanders – lumping the pencil, for instance, and the folder as same colored – are highly unnatural, and he recommends that we let nature determine admissible interpretations. (Let naturally similar things "magnetically" attract one another in the proper extensions.) For Lewis, we must accept only interpretations that conform as much as possible to the way the world actually is. In the famous phrase, a natural interpretation will assign extensions to general terms in a way that "reflects the objective joints in nature."[31] Returning to our case, in order to preserve our categories as unified, we'd have to invoke this "magnetism" for $\mathcal{W}$ and then again for $\mathcal{W}^{\#}$. ($\mathcal{W}$'s magnetic coat would – in virtue of principle # – eliminate some deviance inducing permutations of the domain of $\mathcal{W}^{\#}$; then we'd have apply something like this magnetism to that domain itself in order to rule out the 'not even false' monsters.)

We see, thus, that a semantic grounding of VIFO must be buttressed by the meta-semantic assumption of reference magnetism. Reliably to infer *VIFO* we will need to insulate $\mathcal{W}$ and $\mathcal{W}^{\#}$ with magnetic coats. So now, the purported justification of *VIFO* must go like this: Assume empiricism, do the formal semantics of $\mathcal{W}^{\#}$, and to insure uniformity – then apply two dollops of reference magnetism.

30 Max Cresswell suggested the Whorfian hypothesis.

31 "When we limit ourselves to the eligible interpretations, the ones that reflect the objective joints in nature, there is no longer any guarantee that (almost) any world can satisfy (almost) any theory" (Lewis (1984)).

3.3 The Simchen scramble (non-local truth)

But not so fast. Trouble still lurks. Lewis blocked the gerrymandered extensions that artificially can mar any link to the world from the sentences that we accept as true; and I argued that a similar Lewis-block will protect metaphysical determinations based on true predications. An idea of Ori Simchen's, however, brings back the mischief at the level of truth, specifically at clause T1. Yes, at the point of predication itself. We generally (almost mindlessly) assume that once the references of our singular and general terms are fixed, then truth itself will be transparent. But, says Simchen, it needn't be.

3.3.1 The basic Simchen scramble

His idea: We start with our model $\mathcal{W}$, and once again permute the domain. He calls the function, g, doing that a "scrambler". Next, we let that scrambling go straight for clause T1 and revise it: In general for a model $\mathcal{M}$, we have

$\text{T1}^{!g}$: $\mathcal{M} \models_g Q_k(b)$ iff $g(\mathcal{I}^{\mathcal{M}}(b)) \in \mathcal{I}^{\mathcal{M}}(Q_k)$.[32]

This says that the predication, $Q_k(b)$, is not true in $\mathcal{M}$ in virtue of the referent of b being in the extension assigned by $\mathcal{I}^{\mathcal{M}}$ to Q_k. No, on this conception of truth, $Q_k(b)$ will be true in $\mathcal{M}$ only if the element of the domain that replaces the referent of b under h falls into that extension. Thus, even if reference is kosher, nevertheless, a predication could be true just in case something other than the referent of the name in that predication falls under the extension assigned to the predicate term. This seems obviously a deviant notion of truth, "scrambled truth" Simchen calls it.

Now of course when we allow this scrambling there is no guarantee that the original truths of $\mathcal{M}$ will still come out true. Here's an example: Start with our model $\mathcal{W}$, and let f permute the domain, $D_{\mathcal{W}}$ as above. That is: $f(o_1) = o_3$; $f(o_2) = o_4$; $f(o_3) = o_1$; $f(o_4) = o_2$. But now leave the reference relation for terms untouched, as in Figure 1. Still, under clause $\text{T1}^{!f}$, the sentence "Posy's sweater is blue" comes out false. The reason is that, under this notion of *truth*, that sentence's truth condition is that my pencil be blue, which it is not. So in general when we use this scrambled-truth definition on the model, $\mathcal{W}$, the truths of Σ will not be preserved under $\models_g$.

32 To make this complete, we need to tinker at the level of satisfaction and assignments. This is readily done.

But now the kicker: we can do this scrambling in a way that does preserve the truths of the original model. Here's how:

- First we let the scrambler, say h, permute the domain, $D_{\mathcal{M}}$.
- Then we we redefine reference according to this permutation, as Putnam did.
- Third, we define truth in terms of h^{-1} (the inverse of h).

It is a simple exercise to see that in this case all the truths of $\mathcal{M}$ are formally preserved. Here's an example to make the point:

- Start with our model $\mathcal{W}$.
- Then let's define a model which uses the permutation f of $D_{\mathcal{W}}$ as above; but now once again let f control reference as it did in the Putnamian $\mathcal{W}^{\star}$.
  - So in this model (we can call it $\mathcal{W}^{\star'}$) the expression "Posy's sweater" now refers to my pencil.
- Now let's define truth using clause $T1^{!h}$, but taking g to be f^{-1}.
  - f^{-1} (my pencil) $(= f^{-1}(o_4))$ = my sweater.
  - So the truth condition for "Posy's sweater is blue" is in fact that my sweater be blue, which it is.
- The same will hold for every other elementary sentence of Σ and all the rest as well. All the truths are preserved.

Let me make two comments:

First: Scrambling by itself is not so cockamamie a notion as it might initially seem. It has precedents: Leibniz again is one: truths about empirical objects, for him, rest on truths about monads. But in fact nowadays we've lots of examples of such "non-local" truth: Modal semantics looks beyond the local situation to assign truth-values to compound sentences; so too intuitionistic, temporal, deontic and many more semantic systems. I can even show you free logics with non-local *predication*.[33] Simchen's scrambled truth simply generalizes these.[34] Scrambled truth is a notion of truth.

But, secondly, allow the truth-preserving scrambling and you invite the Putnamian move: you permute the domain to screw things up with one hand, and

33 See Posy (1982).

34 This is what Nino Cocchiarella (1975) called 'secondary semantics'. You could think of a scrambled truth model as a triple $<D_m, I^m, \sigma>$ along the lines of the well-known model structures, where σ is a permutation of the domain, but truth is defined as in $T1^{!}$. Simchen himself suggests that we might judge scrambled truth more sympathetically when we realize that ordinary transparent predication is but a special (limit) case of scrambled truth when σ is the identity function. This argument persuades me less. Identity, it seems to me, is a special notion, not a limit of some other more untamed notion.

then, following Simchen, you compensate with the other. You can mix up objects wildly, but then you can use the new clause $T1^{!h}$ to preserve the truth of your original sentences.

This sort of Σ preserving scrambling is bad enough for those who care about Σ's correspondence to the world. The transparency of predication, as in clause T1, is a crucial link in our inference from what we take as true to how the world is. With scrambling, the sentences of Σ might well be "true" but without telling us anything about the way the world is. But for the profile $\mathcal{W}^{\#}$ – and for our *VIFO* grounding – this is pure disaster. Naturally limit your concept extensions all you like, nevertheless, all our random assignments return at predication. Yet still, once again, we manage to preserve the truths of $\mathcal{W}^{\#}$. This means, once again, that if we maintain the semantic grounding of *VIFO* we cannot assure any reasonable unity to its categories.

3.3.2 A last twist

This is not the end of the mischief. Putnam mixed up the domain and gerrymandered the construction of the extensions that stand in for the properties. Applying this to $\mathcal{W}^{\#}$ is what blocked any claim to unity for the categories of substance and property. However, strictly speaking, what Putnam did is to substitute some elements of the power set of $D_{\mathcal{W}}$ for others. That, in fact, is what the gerrymandered and original extensions amount to: subsets of $D_{\mathcal{W}}$. Recognize this, and you see that there is nothing in principle to stop us from scrambling extensions in the way that we scrambled elements of the domain. There is nothing in principle to stop us from using arbitrary other sets in our scrambled truth conditions rather than the natural extensions of our general terms. If we substitute arbitrary sets for extensions in order to define truth, scrambling can be given by:

$T1^{!!c}$: $\mathcal{W} \models_c Q_k(a)$ iff $\mathcal{I}(a) \in c(\mathcal{I}(Q_k))$, where c is some permutation of $\mathcal{P}(D_{\mathcal{W}})$.

Once we allow this, it turns out that we will mess up the requirement that the *VIFO* categories not overlap. Here's an example to make the point: Let e be a permutation of $D_{\mathcal{W}^{\#}}$ that makes following substitutions:

- $e(\mathcal{I}^{\mathcal{W}^{\#}}(P)) = (\mathcal{I}^{\mathcal{W}^{\#}}(P) - \{o_7\} \cup \{o_2, o_3\})$. (That is $e(\mathcal{I}^{\mathcal{W}^{\#}}(P))$ is the set that is just like $\mathcal{I}^{\mathcal{W}^{\#}}(P)$ but is missing o_7, and has o_2 and o_3 added.)
- $e(\mathcal{I}^{\mathcal{W}^{\#}}(S)) = (\mathcal{I}^{\mathcal{W}^{\#}}(S) - \{o_2\} \cup \{o_5, o_7\})$. (That is $e(\mathcal{I}^{\mathcal{W}^{\#}}(S))$ is the set that is just like $\mathcal{I}^{\mathcal{W}^{\#}}(S)$ but is missing o_2, and has o_5 and o_7 added.)

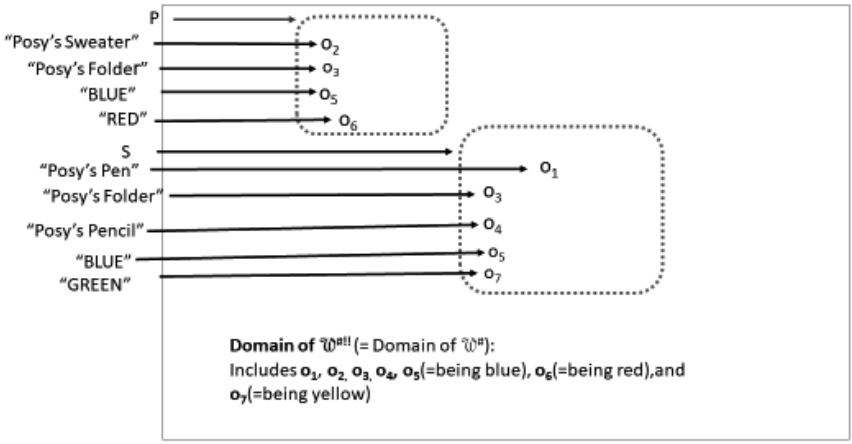
These sets are not disjoint: o_5 and o_3 belong to both of them.

Now we build the following model, $\mathcal{W}^{\#!!}$:

- We start with $\mathcal{W}^{\#}$.
  - And in particular, we leave the reference of names as it is in $\mathcal{W}^{\#}$.
- Next we define $\mathcal{I}^{\mathcal{W}^{\#!!}}(P) = e(\mathcal{I}^{\mathcal{W}^{\#}}(P))$ and $\mathcal{I}^{\mathcal{W}^{\#!!}}(S) = e(\mathcal{I}^{\mathcal{W}^{\#}}(S))$.
- We use principle $T1^{!!c}$ to give the predication and truth conditions in $\mathcal{W}^{\#!!}$, where $c = e^{-1}$.

Here is the story graphically:

Figure 5: **Reference of Terms in the Model $\mathfrak{W}^{\#!!}$**



As I said, in this example, the extensions of *P* and *S* overlap. But using the inverse of the scrambler to define truth (i.e., taking $c = e^{-1}$) in clause $T1^{!!c}$ will still preserve all the truths of $\mathcal{W}^{\#}$. So now, we see that if we insist on a semantic ground for *VIFO* even non-overlap goes by the wayside.

Could we escape this by trying a stronger magnetism, "truth magnetism," in order to rule out such counterexamples? I think not. Such a magnetism would amount to redundancy: saying "ϕ is true just in case ϕ is true." Redundancy conceptions of truth have their advocates and their uses; but for our purposes redundancy is a rather flabby account of truth. It leaves no handle by which we can recursively connect truth to reality. The bottom line: The purported semantic justification of *VIFO* is scotched!

4 A bit of reflection

4.1 Three observations

To see what is going on here, let me first make three observations about the centrality of uninterpreted syntax:

4.1.1 The long arm of syntax

The sharp syntactic distinction between singular and general terms was the foundation of our attempted proof of the unity of *VIFO*'s categories. The notion of a wff rests on this distinction, and this notion in turn dictated without exception the order of the terms in the relation A – the alleged ground for the unity of the categories Property and Substance. Putnam, Quine, Simchen, and Lewis all cleave to this distinction of logical types in regimented languages, like $\mathcal{L}$. However, by letting properties and individuals swim together in $\mathcal{W}^{\#}$'s domain, we wrought the mischief that circumvented that otherwise staid singular-term/general-term distinction. Semantics, if you will, did an end run around the stability of syntax. This ultimately is what allowed us to undermine the claimed unity of *VIFO* categories.

4.1.2 Bridges not joints

The Whorfian (and indeed the Quinean) gloss on those permutation-induced ontologies gives the impression that we are dealing with concept *formation*. Sider's Lewis-inspired talk (in Sider (2011)) of "carving the world at its joints" similarly sounds as if we are *building* semantic correlates of our terms. We are not. Model theory concerns the *correlation* of syntactic terms with extensions (individuals and sets), no matter how these latter may have been singled out. Ultimately, blocking Putnamian mischief, as Lewis wants, is less a matter of carving, and more a matter of creating the proper bridge between two heterogeneous territories: uninterpreted lexical-syntax-land on the one hand and the world of individuals and properties on the other. Natural bridges, not carvings, are the correctives that will give us back our intended truths. Simchen's insight is that the clause for predicative truth – the basis clause in our truth definition – similarly connects reciprocally foreign territories, and opens the door for reference mischief. But, as I pointed out, in this case there appears to be no non-trivial corrective.

4.1.3 The heart of model theory

We need that heterogeneity – in particular we need the isolation of pure, uninterpreted syntax – in order for model theory to do the central task for which it was created: to produce an account of logical truth and logical validity. For, to produce such an account, we need to be able to quantify over different interpretations of a set of sentences. Invariance under different interpretations, that's what gives us logical truth and logical consequence. Model theory's semantic assent and its bridge-notions of reference and of truth serve as a means to define that invariance.

4.2 Not the business of ontology

I mention all of this in order to stress that a theory of logical truth and semantic consequences is not our workshop's promised product. To put it crudely – but not inaccurately – when I'm in the business of determining an ontology I care about my sweater and its properties and not how that sweater fits into a nexus of truths confirmed by the collection of sentences that are true in a model.

This last point echoes the calls of meta-semantic productivism, the view that natural language terms achieve their references "pre-sententially". The whole debate between Putnamian productivism and Lewisian "descriptivism" belongs to the meta-semantic study of how expressions achieve their semantic values. Simchen, for his part, uses his scrambled truth construction mainly in order to support a meta-semantic productivism.[35]

In fact, however, I think that the van Inwagen-style ontological project does not rest on a meta-semantic or even a semantic foundation. Van Inwagen will say that the ontologist need not (and therefore does not) look at the target sentences as initially uninterpreted strings, raw syntax whose role is to be endowed with semantic value. The ontologist simply uses sentences as interpreted. He does not reflect upon them or their relation to the world. He "accedes to his language", in Quine's phrase.

I believe this how we should understand those Van Inwagen remarks about 'Tarskian' and 'satisfiers', remarks so seemingly redolent of model theory. Tarskian is not the metalanguage for some uninterpreted formal language. Rather it itself is ordinary interpreted philosophical discourse with its logical structure exposed

35 See Simchen (2017a,b). Kripke-Putnam causal accounts of reference are the most famous productivist doctrines, but there are other instances as well. The late Putnam's 'direct reference' theory is a prime example. See Putnam (2012b).

and highlighted. It looks directly at its subject matter and not at some uninterpreted linguistic structure.[36] It looks at objects in the world. Yes it may well look at what we say when we speak non-philosophically about the world.[37] But, again this non-philosophical discourse directly depicts the world; not because they are the referents of singular terms, but because they are the subjects of our discourse, what we are talking about. The satisfiers of its open sentences are the objects we take to be in the world. And yes, sometimes indeed we may also look at the structure of that discourse; but then mainly as means of describing relations and determining their 'arity'. In this sense the structure of the discourse is the structure of the world.[38]

Another way to put this is to say that for the ontologist language is a tool, a means of displaying the world; but the world is what interests the ontologist, not the tool, nor how the tool manages to work. The botanist doesn't study the optics of her eyeglasses, and the ontologist does not reflect the abstract theory of reference, at least not *qua* ontologist.

This analogy highlights a stark difference between empirical and metaphysical methods and their respective tools. In an actual physical workshop, tools are robust. Hammers have multiple uses. You can use a hammer to pound in nails, or to push them out, or as a weapon,if you want. You can use eyeglasses as a paperweight. Not so for us philosophers: Our tools must be custom fitted to the job at hand; and if they don't fit they should not be used. Semantics is about the way language connects to its targets. Model theory in particular, aims to explain validity and consequence in discourse – and for that it is fluid about interpretations. But, as we have seen, for that very reason – because of that fluidity – this is ill fitted for our category-fixing task.

36 Thus to complete the quotation from note 16 above. "The vocabulary of Tarskian consists of closed or open sentences and closed or open terms of English (or some natural language) and the sentential connectives, brackets, quantifiers, variables, and identity sign of first order logic (so-called) with identity – perhaps supplemented by items form the vocabulary of various well-defined extension of first-order logic with identity." (Van Inwagen (2014b)).

37 To be sure, van Inwagen (in Van Inwagen (2006)) says that it is unavoidable to focus on the expressions designating non-symmetric relations like "loving" or "being loved by". But, again, what is said and not general truth conditions is at play here.

38 See Van Inwagen (2004).

5 The positive use of model theory

Do I then recommend we ban model theory and its quantification over models from ontology? Absolutely not. For I think there is an ontological task – not category fixing – for which model theory is well suited.

Van Inwagen has the core of the idea: An ontology is a deep and pervasive way of portioning the objects in the world. *VIFO* is one such; but as he is wont to remind us there are other contenders. (He speaks, in particular, of James van Cleve's 'new bundle theory' and Laurie Paul's property-based ontology.[39]) Once we have an ontology – however we may have come upon it, and however we may justify it – we want to endow it with what van Inwagen calls *modal robustness*:[40] It shouldn't matter if I happen to have an older brother (I don't), but if I did, then for van Inwagen, my brother must be an individual substance. Similarly, if there happen to be horses, or centaurs, or unicorns, or even all of these; whatever equine beings there may be, each instance should be an individual substance. By contrast, "being a horse" and "being equine" will be properties, come what may. The point is clear: Objects of either sort may be or not be, but the categories of substance and property (and I would add their exhaustive duality) will be untouched by what actual objects there happen to be, by which model happens to capture the actual world.

The same goes for the rival ontological schemes. Ontologies have this sort of modal force. And now, model theory says this best. There's something inter-model constant about the Substance/Property division. When you advocate *VIFO*, you should be willing to advocate it with this inter-model conviction. In designating these ontological predicates, you should be prepared to accord them that special significance. In our model theoretic terms, in fixing an ontology, you are committed to its holding beyond $\mathcal{W}$.

I am not taking a stand here on some fraught issues about the nature of possible worlds. I would no more identify a possible world with a model – a set theoretic object – than I would identify a property with its extension. On the other hand, just as extensions can partially represent their properties, so too a model may partially and extensionally represent a world. Model theory in this sense gives a concise statement of modal robustness: When you fix on an ontology, you consequently pick out a collection of models.

39 See Van Cleve (1985) and Paul (2002, 2006).

40 "I am inclined to think, therefore, that the account of "ontological category" that I have given needs to be supplemented by a clause to the effect that an ontological category must in some sense be 'modally robust' " (Van Inwagen (2012)).

But then van Inwagen goes a step further. Yes, an ontological category must be modally robust, he says, "but almost certainly not so robust that an ontological category must, by definition, exist in all possible worlds."[41] Taking this at face value, I want to say that this opens a place for the alternative ontologies, those of rival ontologists and even the weird ones generated by Putnam-Simchen scrambling on $\mathcal{W}^{\#}$. Some of these are perhaps nonsense, but they are not ruled out by the model theoretic rules. That fact tells us that whatever ontology one hits upon – and even given its modal robustness – it is in a real sense, contingent. *VIFO* fans have singled out a set of models. This set, however, is but a single galaxy in a space with other ontological galaxies. Each of those galaxies will express the modal robustness of some alternative 'ontology'. I am not saying that any of those other galaxies of models are possible in any real or interesting sense. Most are not; but the collection of them does serve to keep us modally modest: For, our modally robust ontology is, in this second-order sense, contingent!

Model theory can't serve to ground an ontology. It's need (and thus its ability) to quantify over diverse sets of models opens too many possibilities and thus undermines the needed unity and even the distinctness of any purported categories. That we have seen; but now we see that this same quantificational power allows it to create the level of contingency that I have called modal modesty. The categorial distinctions that are invariant within what I called a galaxy of models express the modal robustness of an ontology. That very same ability to group models and compare classes of them now gives the ideal tool to express our modal modesty.

6 A very brief historical allusion

I want to make a quick historical analogy in order to show you why this last point about modal modesty puts model theory back on the shelf of metaphysical tools. The analogy is to Kant's 'transcendental idealism'.

Kant famously holds that space and time – the forms of human perception – constrain the nature of empirical objects. The empirical objects of science and daily life necessarily conform to the principles of mathematics (Euclidian geometry, in particular, for Kant) and to some other spatiotemporally limited conceptual constraints.[42] Leave aside for now his notorious phenomenalist language. If

41 *Ibid*.

42 These latter are what he calls "schematized categories". The Kantian notion of category is, at best, loosely related to van Inwagen's notion of ontological category. "Causality" for Kant is the flagship Kantian category. The abstract notion of causality is the relation of one thing being

you look at Kant's arguments for this doctrine, you will find that they rest on two points: The first is that the realm of what exists and the realm of things that we can empirically recognize and know about, these two realms coincide. His second is that in order for us to know what there is, it must be that geometric and other constraints hold and would hold no matter what empirical objects there might be. He calls this latter doctrine the "*a priori* validity" of geometry and of those other constraints; and the former he calls 'empirical realism'. Well, empirical realism is not far from the 'empiricism' we have been assuming from the start for our $\mathcal{L}$ and its Σ; and '*a priori* validity' exhibits the structure of what we have been calling "modal robustness."

Kant goes on to tell us that we cannot determine the actual content of these *a priori* constraints – the Euclidean axioms in particular, for instance – by any purely logical or analytical methods. The constraints are 'synthetic', he says. Indeed – though this is not often noted – he repeatedly insists that we must consider the abstract notion of beings whose forms of intuition differ from our own; whose 'geometry', for instance, might be non-Euclidean. The way the argument in fact goes is that finite, receptive beings must have a priori perceptual and conceptual constraints in order to have any objective knowledge at all. Thus each species of such beings will have its own modally robust set of *a priori* constraints. Having established that, then we determine from our experience the actual forms that our human perception happens to have.[43]

Kant says quite clearly that we cannot picture the way such beings perceive the world, and he does not claim that such beings or their *a priori* principles are 'really possible' in any graspable way.[44] Moreover, considering them does not detract in any way from Kant's 'emprical realism', his doctrine that what *we* can know is what there is. But, nonetheless we must consider those outlandish beings abstractly in order to fathom the limits of our own *a priori* forms of intuition and thought.[45] This is the heart of his 'transcendental idealism".

---

dependent on another for its existence. The 'schematized' version is expressed by the principle that for each temporal event *A* there is a temporally prior event *B* from whose occurrence *A*'s occurrence follows.

43 Posy (2000, Forthcoming) provide more detail of the Kantian reasoning described here, and put in in historical context.

44 "Other forms of intuition than space and time, other forms of understanding than the discursive forms of thought, or of knowledge through concepts, even if they should be possible, we cannot render in any way conceivable and comprehensible to ourselves. And even assuming that we could do so, they would not belong to experience – the only kind of knowledge in which objects are given to us." (Kant (1929), A230/B283).

45 "Since we cannot treat the special conditions of sensibility as conditions of the possibility of

Nowadays we do not find non-Euclidean geometries to be as outlandish as did Kant; but, nonetheless, I propose to you that Kantian transcendental idealism has the very same structure as our modal modesty.[46] It places the galaxy of spatiotemporal *a priority* in a metaphysical universe of other galaxies, reflecting other (perhaps 'not even false') constraints. It is about the contingency of our human forms of intuition. Kant's doctrine is an idealism because it says that what is 'empirically real' is in fact constrained by the way we humans are built. It is 'transcendental' because – though we discover the constraints empirically – we must nonetheless use philosophical, not empirical, tools to establish them as *a priori* valid.

Kant speaks of idealism; today we speak of anti-realism. I would not presume to attribute such doctrines to van Inwagen; but I do want to suggest that this modal humility is a more palatable anti-realism that Putnam's 1970's version.[47] That, I think, is a worthy product of a metaphysical workshop.

APPENDIX: Technical details of model theory

Grammar of the language $\mathcal{L}$

Vocabulary

G1: There are singular terms:
Names: "Posy's pen", "Posy's sweater", "Posy's folder", "Posy's pencil", $a_1, \ldots, a_n, \ldots$; and
Variables: $x_1, \ldots, x_n, \ldots$;

G2: There are general terms: "BLUE", "RED", "GREEN", and other one place predicate symbols $Q_1, \ldots, Q_n, \ldots$;[48]

things, but only of their appearances, we can indeed say that space comprehends all things that appear to us as external, but not all things in themselves, by whatever subject they are intuited, or whether they be intuited or not. For we cannot judged in regard to the intuitions of other thinking beings, whether they are bound by the same conditions as those which limit our intuition and which for us are universally valid." (Kant (1929), A27/B43).

46 The modal modesty that I am attributing to Kant is different from the "Kantian humility" proposed by Langton (1998). That latter is wholly epistemic and rests on the irreducibility of relations.

47 In Posy (2015) I worked out a version of this suggestion in more detail.

48 I am assuming, recall, that $\mathcal{L}$ is a monadic language.

G3: There are logical symbols: "$\sim$", "$\wedge$", "$\forall$", "(", ")".

Well formed formulas (wffs)

G4: If t is a singular term (name or variable) then for each i, $\ulcorner Q_i(t)\urcorner$ is an atomic wff;

G5: If ϕ is a wff then $\ulcorner \sim \phi\urcorner$ is a wff;

G6: If ϕ and ψ are wffs then $\ulcorner(\phi \wedge \psi)\urcorner$ is a wff;

G7: If ϕ is a wff then for each i, $\ulcorner \forall x_i \phi\urcorner$ is a wff.

Basic formal sematics

Reference

A *model* $\mathcal{M}$, is the pair $<D_{\mathcal{M}}, \mathcal{I}^{\mathcal{M}}>$, where $D_{\mathcal{M}}$ is a non-empty set, and $\mathcal{I}^{\mathcal{M}}$ is a function such that

I1: For each name a_i, $\mathcal{I}^{\mathcal{M}}(a_i) \in D_{\mathcal{M}}$; and

I2: For each general term Q_i, $\mathcal{I}^{\mathcal{M}}(Q_i) \subseteq D_{\mathcal{M}}$.

Satisfaction and truth

Let s be an *assignment* of values to the variables: that is for each i, $s(x_i) \in D_{\mathcal{M}}$.

s induces the following function $\mathcal{I}_s^{\mathcal{M}}$:

- $\mathcal{I}_s^{\mathcal{M}}(a_i) = \mathcal{I}_s^{\mathcal{M}}(a_i)$ for each i;
- $\mathcal{I}_s^{\mathcal{M}}(Q_i) = \mathcal{I}_s^{\mathcal{M}}(Q_i)$ for each i; and
- $\mathcal{I}_s^{\mathcal{M}}(x_i) = s(x_i)$ for each i.

Given a model $\mathcal{M}$ and an assignment s, the *satisfiability relation* $\models_s$ is defined by the following conditions, where t is a term (name or variable), Q is a general term, and ϕ and ψ are formulas:

S1: $\mathcal{M} \models_s Q(t)$ iff $\mathcal{I}_s^{\mathcal{M}}(t) \in \mathcal{I}_s^{\mathcal{M}}(Q)$;

S2: $\mathcal{M} \models_s \ulcorner \sim \phi \urcorner$ iff it is not the case that $\mathcal{M} \models_s \phi$;

S3: $\mathcal{M} \models_s \ulcorner (\phi \wedge \psi) \urcorner$ iff $\mathcal{M} \models_s \phi$ and $\mathcal{M} \models_s \psi$;

S4: $\mathcal{M} \models_s \ulcorner \forall x \phi \urcorner$ iff for every assignment s', which agrees with s except perhaps at x, $\mathcal{M} \models_{s'} \phi$.

The relation $\models$ is defined as follows:

T: $\mathcal{M} \models \phi$ iff for each assignment s, $\mathcal{M} \models_s \phi$

An immediate consequence of these definitions is the principle

T1: For any name a and general term Q, $\mathcal{M} \models Q(a)$ iff $\mathcal{I}^{\mathcal{M}}(a) \in \mathcal{I}^{\mathcal{M}}(Q)$.

Validity

V1: A sentence ϕ is logically valid if for all models $\mathcal{M}$, $\mathcal{M} \models \phi$;

V2: An inference $<\phi_1, ..., \phi_n, \psi>$ is valid if for all models $\mathcal{M}$ such that $\mathcal{M} \models \{\phi_1, ..., \phi_n\}$, M$\models \psi$.

V2 has as its immediate consequence:

V3: ψ is a consequence of a theory axiomatized by a set, Γ, of sentences, if for all models $\mathcal{M}$ such that $\mathcal{M} \models \phi$ for all $\phi \in \Gamma$, $\mathcal{M} \models \psi$.

Bibliography

Cocchiarella, N. (1975), "On the Primary and Secondary Semantics of Logical Necessity", in *Journal of Philosophical Logic*: 4(1), 13–27.
Cocchiarella, N. (1987), *Logical Studies in Early Analytic Philosophy*, Columbus: Ohio State University Press.
Cocchiarella, N. (1996), "Conceptual Realism as a Formal Ontology", in Poli and Simons (1996), 27–60.
Cocchiarella, N. (2001), "Logic and Ontology", in *Axiomatics*: 12, 117–150.
Cocchiarella, N. (2007), *Formal Ontology and Conceptual Realism*, Springer.
Grandy, R. (1976), "Anadic Logic and English", in *Synthese*: 32(3/4), 395–402.
Grandy, R. (1977), *Advanced Logic for Applications*, Synthese Library, vol. 110, Dordrecht: D. Reidel.

Hawthorne, J. and Turner, J. (eds.) (2011), *Philosophical Perspectives 25*: *Metaphysics*, New York: Wiley-Blackwell.
Kant, I. (1929), *Critique of Pure Reason*, translated by N. K. Smith, London: Macmillan and Co.; Subsequent editions published by St. Martin's Press, Inc., New York.
Langton, R. (1998), *Kantian Humility*, Oxford and New York: Oxford University Press.
Lewis, D. (1984), "Putnam's Paradox", in *The Australasian Journal of Philosophy*: 62, 221–236; Reprinted in Lewis (1999).
Lewis, D. (1999), *Papers in Metaphysics and Epistemology*, Cambridge: Cambridge University Press.
Novak, L., Novotny, D., Prokop, S., and Svoboda, D. (eds.) (1998), *Metaphysics: Aristotelian, Scholastic, Analytic*, Frankfurt: Ontos Verlag (in cooperation with Studia Neoaristotelica).
Paul, L. (2002), "Logical Parts", in *Nous*: 36, 578–596; Reprinted in Rea (2008).
Paul, L. (2006), "Coincidence as Overlap", in *Nous*: 40, 623–659.
Poli, R. and Simons, P. (eds.) (1996), *Formal Ontology*, Kluwer Academic Publishers, Nijhoff International Philosophy Series, vol. 53.
Posy, C. J. (1982), "A Free IPC is a Natural Logic: Strong Completeness for some Intuitionistic Free Logics", in *Topoi*: 1, 30–43; Reprinted in *Philosophical Applications of Free Logic*, edited by J. K. Lambert, Oxford: Oxford University Press, 1991.
Posy, C. J. (2000), "Immediacy and the Birth of Reference in Kant: The Case for Space", in Sher and Tieszen (2000).
Posy, C. J. (2015), "Realism, Reference and Reason: Remarks on Putnam and Kant", in *The Philosophy of Hilary Putnam*, vol. XXXIV, edited by R. E. Auxier, D. R. Anderson, L. E. Hahn, Open Court, 565–598.
Posy, C. J. (Forthcoming), "Of Griffins and Horses: Mathematics and Metaphysics in Kant's Critical Turn", in Posy and Rechter (Forthcoming).
Posy, C. J. and Rechter O. (Forthcoming), *Kant's Philosophy of Mathematics*, Volume 1: The Critical Philosophy and Its Roots, Cambridge: Cambridge University Press.
Putnam, H. (1977a), "Realism and Reason", in *Proceedings and Addresses of the American Philosophical Association*: 50(6), 483–498; Reprinted in Putnam (1978).
Putnam, H. (1977b), "Models and Reality", in *Presidential Address to the Winter Meeting of the Association for Symbolic Logic*, Washington D. C.; Reprinted in *Journal of Symbolic Logic*: 45(3) (1980), 464–482; and in Putnam (1983).
Putnam, H. (1978), *Meaning and the Moral Sciences*, Oxford and Boston: Routledge and Kegan Paul.
Putnam, H. (1983), *Realism and Reason: Philosophical Papers*, vol. 3, Cambridge and New York: Cambridge University Press.
Putnam, H. (2012a), *Philosophy in an Age of Science*, Cambridge MA: Harvard University Press.
Putnam, H. (2012b), "Corresponding with Reality", in Putnam (2012a).
Rea, M. C. (ed.) (2008), *Critical Concepts in Philosophy*, vol. *Metaphysics*, London and New York: Routledge.
Russell, B. (1918), *The Philosophy of Logical Atomism*, reprinted in 1985, London and New York: Routledge.
Russell, B. (1937), *A Critical Exposition of the Philosophy of Leibniz*, London: George Allen and Unwinn, 2nd Edition.
Russell, B. (1938), *Principles of Mathematics*, New York: W. W. Norton, 2nd Edition.
Sher, G. and Tieszen, R. (eds.) (2000), *Between Logic and Intuition*: *Essays in Honor of Charles Parsons*, Cambridge and New York: Cambridge University Press.

Sider, T. (2011), *Writing the Book of the World*, Oxford and New York: Oxford University Press.
Simchen, O. (2017a), "Metasemantics and Singular Reference", in *Nous*: 51(1), 175–195.
Simchen, O. (2017b), *Semantics, Metsemantics, Aboutness*, Oxford: Oxford University Press.
Strawson, P. F. (1974), *Subject and Predicate in Logic and Philosophy*, Ashgate Publishing, 2nd Edition, 2004.
Van Cleve, J. (1985), "Three Versions of the Bundle Theory", in *Philosophical Studies*: 47, 95–107.
Van Inwagen, P. (2004), "A Theory of Properties", in Zimmerman (2014), 107–138; Reprinted in Van Inwagen (2014a).
Van Inwagen, P. (2006), "Names for Relation", in *Philosophical Perspectives*: 20, 435–477.
Van Inwagen, P. (2011), "Relational vs. Constitutive Ontologies", in Hawthorne and Turner (2011), 389–405; in Van Inwagen (2014a).
Van Inwagen, P. (2012), "What is an Ontological Category", in Novak *et al*. (1998), 11–24; in Van Inwagen (2014a).
Van Inwagen, P. (2014a), *Existence*: *Essays in Ontology*, Cambridge: Cambridge University Press.
Van Inwagen, P. (2014b), "Inside the Ontology Room", in Van Inwagen (2014a), 1–4.
Zimmerman, D. (ed.) (2014), *Oxford Studies in Metaphysics*, Oxford: Oxford University Press.

Part III: Existence, Nonexistence, and Contradiction

Friederike Moltmann

Existence Predicates

Abstract: The standard view about existence is that existence is a univocal concept conveyed by the existential quantifier. A less common philosophical view is that existence is a first-order property distinguishing between 'nonexistent' and existing objects. An even less common philosophical view is that existence divides into different 'modes of being' for different kinds of entities. It is a fact that natural languages generally distinguish among different existence predicates for different types of entities, such as English 'exist', 'occur', and 'obtain'. The paper gives an in-depth discussion and analysis of a range of existence predicates in natural language within the general project of descriptive metaphysics, or more specifically 'natural language ontology'.

Existence is a central notion in metaphysics and it is associated with three important questions:

1. Is existence a univocal concept or are there different modes or degrees of being, different forms of reality, for different types of entities?
2. Are there entities that have 'being' in some sense but not existence?
3. Is existence conveyed by predicates such as *exist* or by quantifiers?

This paper addresses these questions from the point of view of natural language. Thus, it addresses the question whether natural language reflects a distinction between different modes of being or whether it displays a univocal notion of existence. It addresses the question whether natural language involves an ontology of entities that have being but not existence. And it examines the way existence is conveyed, by existential constructions or by existence predicates such as *exist*. Focusing on English, this paper argues for the following answers to these questions. First, natural language displays particular kinds of modes or ways of being, but conveyed by different existence predicates such as English *exist*, *happen*, and *obtain*. Second, natural language reflects a Meinongian view, with an ontology that includes entities that have being but not existence. Finally, existence in natural language is semantically conveyed by existence predicates, not existential quantifiers. Quantifiers, at least in English, are neutral as regards different modes of being associated with different sorts of entities as well as existent and nonexistent entities. Existence predicates, at least in English and related languages, convey particular ways of being, which are not the modes of being generally distinguished in contemporary or historical philosophy. English existence predicates convey ways for entities to relate to time (persistence conditions), and more generally space and time. Thus, *exist* has both a time-related use and, as

https://doi.org/10.1515/9783110664812-009

pointed out by Fine (2006), a space-relative use, both of which impose particular restrictions on the types of entities they can apply to. The paper will explain those restrictions and argue that the time-independent use of *exist* is derivative upon the time-dependent use.

This paper focuses almost entirely on existence predicates in English, *exist*, *occur*, *happen*, and *obtain*. This is in accordance with the general view in contemporary linguistic theory, that an in-depth investigation of a relevant phenomenon in a particular language generally is likely to reveal something fundamental to all languages or our cognitive system as such. This is the established view in generative syntax, but there are equally good reasons to uphold it for semantics as well as for natural language ontology, at least given a suitable range of linguistic data.[1]

This paper is not a paper in linguistic semantics, though, but rather in descriptive metaphysics, more specifically natural language ontology. This means that the presentation and discussion of the linguistic data will throughout be pursued in relation to relevant views in metaphysics. The paper will first situate its project within the context of contemporary metaphysics. It will then discuss quantification and reference in natural language in regard to the notions of existence and modes of being and then come to the main part of this paper: existence predicates in English with their time and space-related uses.

1 The background: Natural language ontology, descriptive metaphysics and foundational metaphysics

For some philosophers the focus of this paper on natural language may raise the question, why should the way natural language reflects existence matter for the philosophical debate? There is an inclination among contemporary philosophers to think that whatever natural language may display, it will be independent of the question of what existence really is, whether there really are things that have being but not existence and whether there are really different modes of being. It is therefore important to situate the question about natural language properly and give it its legitimacy within the philosophical context.

1 The restriction to a particular language is also in accordance with the practice of descriptive metaphysics as such. The linguistic data and generalizations, even if just from a particular language, are a manifestation of the sorts of fundamental intuitions that descriptive metaphysics aims to make use of.

Two different branches of metaphysics should be distinguished. One is what Fine (2017) calls *foundational metaphysics*. Foundational metaphysics asks the question of what there really is; its subject matter is the general nature of reality. Foundational metaphysics thus addresses the question of what existence really is, whether there are really different modes of being and whether nonexistent objects are really there. The other branch of metaphysics is what one may call, following Strawson (1959), *descriptive metaphysics*.[2] Descriptive metaphysics has as its subject matter the ontology reflected in our ordinary judgments. This is not the metaphysics of 'what there really is', but of 'what there appears to be', the 'metaphysics of appearances', as Fine (2017) calls it. A particular version of descriptive metaphysics is the metaphysics that focuses on linguistically reflected intuitions, *natural language ontology* (Moltmann (2017a)).[3] Natural language ontology aims to uncover the metaphysics that is implicit in natural language and that may diverge both from the reflective metaphysics of speakers of the language (or that particular philosophers may be willing to accept) and the metaphysics of what there really is. For example, natural language permits reference to a great wealth of derivative or minor objects (nonworldly facts, institutional roles, shadows etc) many of which many speakers upon reflection may not be willing to accept as objects of their own. Even though it does not directly concern the nature of reality, the metaphysics implicit in our way of speaking or thinking is clearly an important subject matter in itself. The metaphysics reflected in natural language certainly belongs to our cognitive faculty just as much as the syntax and semantics of natural language do.

Also the notion of existence is reflected differently in natural language than in the way various philosophers have conceived of it. While it is an interesting question why that should be so, this paper will rest with the descriptive aim of establishing a range of generalizations about the reflection of existence in English and in particular the notions of existence that English existence predicates convey. Here just a few words regarding common conceptions of existence in philosophy.

The standard view about existence in contemporary metaphysics, mainly due to Quine, is that existence is a univocal concept conveyed by the existential quantifier (in a logical language). A less common philosophical view is that existence is a first-order property distinguishing between nonexistent (past, possible, or merely intentional) objects and existing objects. Such a view generally draws a sharp distinction between the existence predicate *exist*, which conveys existence

2 Fine (2017) call this 'naive metaphysics'. However, 'descriptive metaphysics' certainly is the better established term.

3 See also Bach (1986), who uses the term 'natural language metaphysics'.

in that sense, and the existential quantifier, which is neutral as regards existent and nonexistent objects.[4] An even less common philosophical view in the history of analytic philosophy is that existence divides into different 'modes of being' for different kinds of entities, that is, the view that different kinds of entities have different kinds of reality.[5] The latter view has received some renewed attention, in the form of ontological pluralism.[6] This more recent interest in modes of being is firmly situated within foundational metaphysics and focuses on distinctions between fundamental or natural entities and less fundamental or natural ones. More fundamental entities have a greater 'degree of being' than less fundamental ones. Thus, entities such as holes and shadows have a lesser degree of being than ontologically independent ones.[7] In the recent literature, modes of being in that sense have been tied to different quantifiers, rather than different existence predicates McDaniel (2010a,b)).[8]

We will see that natural language (or at least English and related languages) does not reflect different modes of being based on fundamentality, but rather based on how entities relate to space and time.[9] Moreover, natural language (or at least English and related languages) does not reflect modes of being with different quantifiers, but rather with different existence predicates.

4 See, for example, Salmon (1987, 1998), Zalta (1983, 1988), Muyskens (1989), and Priest (2005).

5 See McDaniel (2009) for a recent discussion of the view of existence dividing into different modes of being.

6 See Turner (2010), McDaniel (2010a,b, 2013).

7 For a recent defense of the Quinean view see Van Inwagen (1998). The view of existence dividing into different modes of being had been held by Aristotle, Heidegger, Sartre, and Moore.

8 An older view about modes of being, involves a distinction between the being of material objects and the 'existence' of free agents (and perhaps the transcendence of god). This distinction is associated with Augustin, Sartre, as well as Jaspers.

9 More generally, it appears that what is fundamental will receive little elucidation from natural language. Natural language, for example, does not distinguish between natural and non-natural properties, but displays a wealth of predicates (and their corresponding nominalizations) expressing non-natural or abundant properties.

2 Quantification and the expression of existence

Let us start with question 3, how existence is expressed on natural language, and in that regard also address question 2, whether natural language permits reference to intentional (nonexistent) objects.

The common, Quinean view is that existence is expressed by quantification or the *there is*/*are*–construction. In natural language, however, existential quantifiers and *there is*/*are* do not as such convey existence. Natural language rather reflects the Meinongian view according to which quantifiers such as *a*, *some*, *two* and *there is*/*are* are neutral regarding existence and non-existence, as is the use of 'referential' singular terms (names and definite descriptions).[10]

They can all be used to talk about 'nonexistent' entities.[11] In natural language, existence is not expressed by quantifiers, but instead by existence predicates such as *exist* in English. This is reflected, for example, in the possible truth of Meinongian statements such as (1), where *there are* ranges over things of which existence is denied:

(1) There are objects of thought that do not exist.

Quantifiers at least in English and other European languages do not distinguish among different modes of being for different sorts of entities. Rather, the very same quantifiers are used to quantify over entities of any sort as well as over past and intentional objects. If this is a crosslinguistic generalization, then this means that a view such as McDaniel (2010a,b, 2013), which posits different quantifiers for different modes of being, is not reflected in natural language.

Not only quantifiers, but also singular terms in natural language are neutral as regards existence. This is most obvious when singular terms occur in the subject position of a negative existential, as below:[12]

10 Meinongians have long argued that existential quantification, unlike predication with *exist*, is not existentially committing. Recent Meinongians include Parsons (1980), Priest (2005) and Zalta (1983, 1988), as well Salmon (1987, 1998) (for past and possible objects only).

11 Note that existential quantifiers and *there is*/*are* can be *used* to express existence, as below:

(i) a. There aren't any objects of thought.
b. There aren't any objects like tables and chairs.

12 There is a common view, held most notably by Frege, that *exist* is a special, second-order predicate, applying to concepts rather than individuals (and thus expressing instantiation or non-emptiness of extension). From the point of view of natural language this is not plausible, though.

(2) a. The king of France does not exist.
b. Vulcan does not exist.

In negative existentials, the subject term, on one view, is an empty term, *exist* expresses the trivial property everything has, and negation is taken to be external negation (so that (2b) roughly means 'it is not true that Vulcan exists').[13] On another view, the Meinongian view, the subject in a true negative existential always stands for an entity, but a 'nonexistent' one, an entity of which an existence predicate such as *exist* is false.[14] There is considerable support for the Meinongian view from natural language. This support, though, does not so much come from sentences that the philosophical literature has focused on, namely sentences such as (1) and (2a, b). (1) is a quasi-philosophical statement, and as such not truly indicative of the ontology implicit in natural language.[15] For negative existentials with simple definite NPs and names in subject position such as (2a, b) there are alternative analyses on which they do not involve NPs standing for objects at all.[16] The support for the Meinongian view being reflected in natural language rather comes from constructions whose *compositional* semantics *requires* intentional objects as semantic values, constructions which are usable without philosophical reflection. These are noun phrases modified by relative clauses with first intensional predicates (in a broad sense, including temporal predicates, which shift

Exist does not require, like putative second-order predicates, predicative expressions. Rather it requires expressions in subject position that act as singular terms, as in (4a, b) and (5a,b), or that act as quantifiers binding individual variables, as in (4c). Further support for the status of the subject of *exist* as a singular term or quantifier comes from its support of anaphora, which are not anaphora relating to predicates as antecedents (*one*, *that*), but anaphora relating to singular terms (*he*, *she*, *it*).

Also philosophers have argued for *exist* not being a second-order predicate, but a first-order predicate, including Miller (1975, 1986, 2002), Salmon (1987, 1998), and McGinn (2000).

13 See Sainsbury (2005) as a representative of that view.

14 There is also a third, hybrid view, that of Salmon (1987, 1998), on which the subject term in true negative existentials sometimes stands for an object of which *exist* is false, namely an object that has existed only in the past or a merely possible object. If the subject of the negative existential is a fictional term, though, Salmon takes it to be empty, with negation then being external negation.

15 For the view that quasi-philosophical statements, statements that imply a certain amount of philosophical reflection, should not be taken into account for the ontology reflected in natural language see Moltmann (2017a).

16 Thus, the Russellian view does not as such take definite NPs to be singular terms, and names in negative existentials have been treated as empty names (see Fn 13). See also Azzouni (2010) for a non-Meinongian account of simple negative existentials.

the time of evaluation) and second intentional predicates, predicates like *read about, hear about, describe, mention*, and *think of*:[17]

(3) a. There are some buildings that were built in the last century that no longer exist.
b. There are some events that John read about / heard about that did not take place.
c. There are buildings described in the book that do not exist.

(4) a. Some buildings the guide mentions do not exist.
b. Some things John thought of do not exist, for example Vulcan.

Sentences like (3a, b, c) and (4a, b) are part of ordinary discourse and do not presuppose any form of philosophical reflection. Also descriptions with intensional or intentional predicates require their semantic value to be a 'nonexistent' object:

(5) The building built there last year no longer exists.

(6) a. The book John is thinking about does not exist.
b. The building mentioned in the guide does not exist.

In (5), past tense allows a nonexistent object to act as semantic values of the descriptions in subject position. In (6a, b), the subject consists of a definite description formed, crucially, with an intentional verb, such as *mention* or *think of*. Such verbs take intentional 'nonexistent' objects as arguments when the intentional act they describe is not successful, and these entities should be the ones the existential quantifiers in (6a, b) range over.[18]

If intentional 'nonexistent' objects are involved in existentially quantified negative existentials as in (6), then they may just as well be part of the semantics of negative existentials with 'empty' proper names or descriptions associated with a failed or pretend act of reference.

There are, of course, also positive existentials:

(7) a. The president of France exists.
b. Mars exists.

Exist as a first-order predicate in positive existentials as in (7a, b) is usually taken to express a trivial or almost trivial property, the property every entity has or, less trivially, the property that every present and actual entity has. However, posit-

17 See the discussion in Moltmann (2015), where the notion of an intentional predicate and the difference between intensional and intentional predicates are discussed in greater detail.

18 See McGinn (2000) for a philosophical defense of that view, and Van Inwagen (2008) for a critical discussion.

ive existence statements do not generally express a trivial truth, and we will see that the primary use of *exist* is in fact a time- or location-dependent use, which does not lead to sentences that are trivially true. Time-relative uses of *exist* permit highly informative statements about past and future existents.

3 Existence predicates in natural language

Existence in English thus is expressed by existence predicates. But what defines a predicates as an existence predicate? What characterizes *exist* as an existence predicate is its behavior in negative existentials as in (1-6). Existence predicates differ from ordinary predicates in that with a non-referring subject (or rather a subject term not standing for an actual object), they generally yield true sentences if they are in the scope of negation, as in (1-6), and false sentences if they are not in the scope of negation.

Exist is not the only existence predicate in English. In fact, natural languages generally do not display a single existence predicate, but different existence predicates for different types of entities. Such selectional restrictions are imposed whether or not the entities to which the existence predicates apply 'exist'. In English, at least, the restrictions of existence predicates to particular types of entities are linked to the fact that those predicates have time- and space-relative uses besides an absolute one, as we will see.

The philosophical concept of existence is generally taken to apply to any actual entity whatever its type. However, the English predicate *exist* in fact applies only to certain types of entities and that not only for 'ordinary' speakers (non-philosophers), but also philosophers when engaging in ordinary discourse.[19] Roughly, the generalization is that *exist* applies to material and abstract objects and is inapplicable to events (Hacker (1982a), Cresswell (1986)):[20]

19 It is remarkable that philosophers even if they have a reflective notion according to which existence is univocal are unable to use *exist* for events. Interestingly, the nominalization *existence* appears able to convey the unvocal concept of existence, as well as covering different modes of being at once, depending on the philosophical view of the language user. Nominalizations thus may convey a notion of a speaker's reflective metaphysics, but not the verb, which is restricted to notion of the metaphysics implicit in language. The structure of language thus appears to along with different degrees of implicitness of the metaphysics adopted by speakers.

20 *Exist* is also not particularly good when applied to a person (with the time-related use), an observation I will set aside in the context of this paper:

(i) ? John's child still exists.

(8) a. The book exists.
b. The round square does not exist.
c. ??? The party/ demonstration exists.
d. ??? The accident John mentioned did not exist.

Here and throughout this paper '??' means 'unacceptable semantically' (though not ungrammatical, that is, syntactically incorrect).

There are specific existence predicates in English for events, namely *occur*, *happen* and *take place*:[21]

(9) a. John's party is taking place.
b. The accident John mentioned occurred / happened.

(10) a. The murder occurred / happened / ??? existed yesterday.
b. John's speech took place / ??? existed this morning.

The test for existence predicates applies straightforwardly to *occur*, *happen* and *take place*. That is, with negation (9a, b) can still be true.

Two further existence predicates in English are *obtain* and *hold*. They apply to 'condition'-like entities, as I will call them, such as states, situations, conditions, rules, laws, and non-worldly facts:[22]

(11) a. The situation /state / condition / law / rule Mary describes (no longer) obtains / holds / exists.
b. The fact that S does obtain / hold.

Also *exist* may apply to some of the conditions-like entities (states, conditions, laws). *Obtain* and *hold* are not applicable to material objects, persons, and abstract objects of the sort of mathematical objects.

Yet another predicate in English that *ca*n be used as an existence predicate is *is valid*. (*Is valid* has other uses as well, of course, for example when applying to arguments or syllogisms). When used as an existence predicate, *is valid* applies only to condition-like objects, just as *obtain* and *hold*, but *is valid* is further restricted to normative condition-like objects like laws and norms. When used as an

21 There are further restrictions on event-related existence predicates. *Occur* goes well with incidents and processes, but not with activities. *Take place* imposes its own constraints on a prior planning of an action, see Section 8.

22 Interestingly (non-worldly) facts do not go with *exist*, but only with *obtain* and *hold*. For the notion of a non-worldly fact see Strawson (1950). Non-worldly facts differ from worldly facts in the sense of Austin (1979), fully specific facts that are part of the world and are rather event-like. See also Moltmann (2013a,b), namely, for the notion of a nonworldly fact (and nonworldly state).

existence predicate applying to normative condition-like objects, it is generally interchangeable with *exist*:[23]

(12) a. The law is valid / exists.
b. That law is no longer valid / no longer exists.

The possible truth of (12b) makes clear the status of the use of *is valid* as an existence predicate, on a par with *exist*.

Given the fact that existence predicates generally impose type restrictions, the characterization of existence predicates given at the beginning of the last section requires a modification. With subject term not standing for an actual object but meeting the type restrictions, an existence predicate generally yields a true sentence if it is in the scope of negation and a false sentence if it is not in the scope of negation. Below the criteria distinguishing ordinary predicates (i.e. predicates that are not existences predicate) and existence predicates are given more formally:

(13) a. A (intransitive) predicate *P* is an *ordinary predicate* iff for any world *w* and time *t*, for any singular term *T*, if *T* does not stand for an actual entity in *w*, then neither $[T\ not\ P]^{w,t}$ = true nor $[T\ not\ P]^{w,t}$ = false.
b. An (intransitive) predicate P is an *existence predicate* iff for any world *w* and time *t*, for any singular term *T*, if *T* satisfies the selectional restrictions of *P* and does not stand for a (present, actual, nonintentional) entity in *w*, then $[T\ not\ P]^{w,t}$ = true and $[T\ P]^{w,t}$ = false.

(13b) rules out a range of predicates that allow for terms not standing for actual objects, but are not existence predicates. Thus, some predicates such as *is important*, *has influence*, *is a philosopher* can apply to past objects. But applied to past or future objects, they do not generally yield true sentences with negation and false sentences without negation. Similarly, predicates like *think about*, *plan*, and *imagine* can apply to apparently empty terms, but they too do not generally yield true sentences with negation and false sentences without negation.[24]

Given (13b), moreover, the predicates *live* and *be alive*, which one might consider existence predicates, do not come out as existence predicates, unlike *exist*:

23 The various existence predicates in English raise the question of how many different types of existence predicates there are in natural languages in general. This is a question highly worth a crosslinguistic study, but pursuing it goes far beyond the scope of this paper, which restricts itself to the way the notion of existence is reflected in English (and related languages).

24 Note that such predicates do not come out as ordinary predicates, which can apply only to presently existing entities. However, this restriction could easily be changed if so desired, allowing for a larger class of ordinary predicates.

(14) a. ??? The president of France does not live / is not alive.
b. The president of France does not exist.

Live and *is alive* presuppose that the object they apply relative to a time *t* was alive before *t*, and thus do generally yield true sentences with negation and a term not referring to an actual entity.

Occur is used both as an existence predicate, as in (15a), and as an ordinary predicate, as in (15b, c):

(15) a. The accident John mentioned did not occur.
b. The letter occurs twice in the sentence.
c. The letter does not occur in the sentence.

Occur is used as an existence predicate in (15a), but not in (15b) and (15c), according to (13b).

The criterion in (13b) also rules out as existence predicates some other predicates that according to particular historical philosophical views might be regarded as such, for example, in phenomenology, *being experienced*.

4 Existence predicates for objects and for events

We can now turn to the semantics of existence predicates, in English. Let us start with the two existence predicates *exist* and *occur*. *Exist*, recall, applies to enduring objects as well as abstract objects, whereas *occur*, as an existence predicate, applies to events only. An important observation is that *exist* and *occur* display those type restrictions both on a time-relative interpretation and when used time-independently. On its time-relative use *exist* conveys endurance. Since it is subject to the same type restrictions, the meaning of *exist* when used time-independently must be derivative upon the meaning of *exist* when used relative to a time, thereby imposing the very same type restrictons.

In this paper, I will make use of a common, though not uncontroversial view of endurantism based on the notion of complete presence. Given the present purpose, I will not give an in-depth discussion of that view and its controversies.[25] The subject matter of thise paper is the metaphysical notions that are reflected in existence predicates in natural language, not metaphysical notions of persistence

25 See, for example, Sider (2001) and Hawley (2001) for discussion.

as such. That is, the paper is a study within natural language ontology, not within metaphysics as such.[26]

The standard formulation of the endurantist view about persistence of material objects through time involves the notion of complete presence (Wiggins (1980), Lewis (1986)): An object exists at a time interval *t* just in case is *wholly present* at each moment of *t*. That is, existence at a time *t* means *complete presence* at all the moments of *t*. A notorious problem with this formulation concerns the condition of 'complete presence'. The standard condition applies only to entities of the sort of sums, whose (non-interchangeable) parts need to strictly be present at each moment of a sum's lifespan. But clearly, for ordinary material objects, in general not all material parts need to be present at each moment of the object's life span. Since this paper will not engage in an extensive discussion of endurantism as such, it will simply subscribe to a suitable weakening of the condition on 'complete presence', to the effect that not for all sorts of objects all material parts need to be present at any moment of their lifespan. For organisms and artifacts, for example, it should be enough that just sufficiently many 'functional parts' be instantiated throughout the time in question. What may also have to be present, moreover, are essential features or the 'way' of composition. For example, objects like statues or stones need to retain more or less their shape throughout the time at which they exist.

Endurantism, on roughly the understanding above, goes along with three-dimensionalism, the view that objects cannot have temporal parts, but only spatial parts, whereas events can have temporal parts. Three-dimensionalism contrasts with four-dimensionalism, the view that objects and events are both space-time regions with spatio-temporal subregions as parts (Sider (2001)).

Three-dimensionalism makes a sharp distinction between the *existence* of an entity at a time (or endurance throughout the time) and its *extension* at a time, which roughly means the locatedness of the entity at that time (or its perdurance throughout the time) (Fine (2006)). Analogously, there are the notions of existence at a space and extension (locatedness) at a space (Fine (2006)). Material objects *exist* in time and are *extended* in space, whereas events are *extended* in both space and time and do not *exist* in time or space.[27] The notions of existence (at a loca-

26 That said, it may not be excluded that the semantics the paper gives for existence predicates in English may, to an extent, be reformulated on the basis of somewhat different philosophical notions or views.

27 Unlike for material objects, the spatial location of events is notoriously difficult to specify. Neither the location of the event participants at the relevant time nor the parts of the participating objects affected by the change induced by the event guarantee an intuitively clear notion of the location of an event. There is in fact another view, defended by Hacker (1982b), on which

tion) and extension are intuitive notions, which are in part reflected in language, though require some clarifying remarks. Existence (at a location), of course, is what is conveyed by location-relative *exist*. Fine (2006) takes extension to be expressed by the combination *be*+location modifier. As Fine notes, *be*+spatial modifier applies to objects, but not *be*+temporal modifier:

(16) a. John was in the garden.
b. * The car was last year.

Thus, the construction captures correctly the spatial extendedness of material objects as well as their failure to be temporally extended. Note, though, that neither *be*+spatial modifier nor *be*+temporal modifier is in fact acceptable with events:

(17) a. ?? John's walking was in the garden.
b. ?? The incident was in Jean's office.

(18) a. ?? John's walking was yesterday.
b. ?? The incident was last year.

events simply do not occupy space and thus cannot be extended in space. If events are attributed a spatial location, then that location would be derivative upon the spatial location of the event participants. The view that events lack a spatial location may be particularly plausible if events are transitions among tropes, that is, particularized properties. While on a common view (following Williams (1953)), tropes come with a spatio-temporal location – or rather with relations of spatio-temporal co-location, tropes in fact do not easily allow for the kinds of location modifiers that one would expect on that view. Thus *be*+spatial modifier is generally excluded with tropes:

(i) a. ?? The apple's greenness was on the table.
b. ?? The roundness of the ball was in the basket.

Moreover, trope-referring terms disallow spatial adnominal modifiers (on the 'normal' interpretation of a spatial modifier):

(ii) a. * the apple's roundness on the table
b. * John's heaviness in the bed

This suggests that it is only the bearer of a trope, not the trope itself that has a spatial location. Tropes depend on a bearer; but unlike their bearer they could not be spatio-temporally located. A trope may ontologically depend on another trope of extendedness, as Husserl had argued (Simons (1994)). But a trope of extendedness is not itself extended; it only instantiates extension.

Be+location modifier is acceptable only when the events are 'movable' events (Dretske (1967)), that is, events for which their spatial or temporal location is not essential, as in (19) and (20):

(19) a. The meeting was in this room.
b. The meeting was moved from one room to another.

(20) a. The meeting was yesterday.
b. The meeting was moved from Monday to Wednesday.

The common view about events is that events have their spatial and temporal location essentially (in particular if events are viewed as instantiations of properties in locations), with the exception, one would have to add, for events that a movable, such as meetings.[28] *Be* + spatial / temporal modifier then appears to presuppose that the entity it applies to does not have its location essentially. Thus, *be* + spatial / temporal modifier appears to convey accidental extension, not extension as such. For an entity's (accidental or essential) extension at a location, I will therefore rather use 'is present at'.[29] The perdurance of an event e throughout a time interval t then requires that for each subinterval t' of t, there is a temporal part e' of e that is present at t'.

The lexical meanings of time-relative *exist* and *occur* in first approximation can now be given as follows:

(21) a. For an entity d (with more than one part), *exist* is true of d at a time t iff for any moment t' of t, (the whole of) d is present at t'.
b. For an entity e (with more than one part), *occur* is true of e at a time t iff for any proper part t' of t, there is a proper part e' of e present at t' and for any proper part e' of e, there is a proper part t' of t such that e' is present at t'.

(21a) requires an entity to which *exist* applies to have more than one part. (Note that this does not prevent it from being ato, which thus does not apply to instantaneous events. (But 'instantaneous' events actually involve at least two mo-

28 See Casati and Varzi (2005) for discussion.

29 One might take time- or space relative *occur*, *happen*, and *take place* to serve that function. But it is not plausible that *occur*, *happen*, and *take place* express extension: *occur* is an eventive predicate, not a stative predicate like *be*+location modifier, and thus it could not express the notion of extension.
Note that *occur*, *happen*, and *take place* unproblematic with spatial or temporal modifiers (*the walking took place in the garden, the incident happened in John's office*), but that is because the location modifier here may give the essential location of an event (of occurrence), unlike when the location modifier is the complement of *be*.

ments, the moment before and the moment after the change of which the event consists.)

(21a) and (21b) do not yet account for the actions art of *exist* and *occur*, that is, the fact that *exist* is a stative verb, whereas *occur* is an eventive verb. Nominalizations such as *the existence of the building* describe states, whereas nominalizations of the sort *the occurrence of the protest* describe events. The latter can have typical event properties such as being sudden or being quick; the former cannot. Adopting a Davidsonian semantics of events according to which verbs take events (and states) as implicit arguments, this means that *exist* takes states ('existences') and *occur* takes events ('occurrences') as implicit arguments. *Exist* when applied to an object d then describes a state that is the presence of 'the whole' of d during the time in question, as roughly given in (22a). *Occur* applied to an event e describes another event that consists in the transitions among the 'presences' of the parts of e at relevant subintervals that belong to the duration of e, as roughly in (22b):

(22) a. For an entity d (with more than one part), an event e, and a time interval t, $< e, d >\in [exist]^t$ iff e consists of the presence of (the whole of) d at t' for each moment t' of t.
b. For an entity e' (with at least two parts), an event e, and a time interval t, $< e', e >\in [occur]^t$ iff e consists of transitions from the presence of e'' at t' to the presence of e''' at t'' for any minimal temporal parts e'' and e''' of e' for which t' and t'' of t at which e'' and e''' are present.

The selectional restrictions of time-relative *exist* and *occur* are presuppositions of (22a, b). (22a) is applicable only to entities that do not have temporal parts, yet are in time, such as material objects; (22b) is applicable only to entities that have temporal parts, that is, events.

Fine (2006) gives a different condition on endurance or time-relative existence, which avoids making reference to parts of an entity. Fine distinguishes extension (in basically the sense above) and location-relative existence in terms of conditions on sums of entities: the sum of two entities is extended at a location l just in case one of the two entities is at l, and the sum of two entities exists at a location l just in case both entities are at l. I will not discuss this view in further detail. One the reason why I will not make use of it is that it does not permit an application of space-relative existence to the class of entities that in fact engage in it, such as languages, illnesses and kinds (Section 5) as well as condition-like entities such as states, laws, and rules (Section 6). Those entities are not the sums of things that need to all be present at the space at which they exist; rather they are composed of abstract parts that need to be instantiated or in place throughout the space at which they exist (Section 5, 6).

5 Space-relative existence

Exist and *occur* exhibit particular constraints when relativized to time or space. The generalizations about time-relative and space-relative interpretations of *exist* and *occur* include the following. Time-relative exist applies to material objects but not to events, whereas space-relative *exist* applies to neither. *Occur* applies to events, both with temporal and spatial modifiers, but not to material objects:

(23) a. The house existed till last year.
b. ?? The incident existed two years ago.

(24) a. ?? The house exists in Munich.[30]
b. ?? The incident existed in this room.

(25) a. ?? The house occurred last year.
b. The incident occurred yesterday.
c. The incident occurred right here.

A brief comment is needed regarding the expression of time-relative and space-relative existence. *Exist* as an English verb always comes with tense morphology, which does not mean, though, that it is always interpreted relative to a time. Present tense only optionally goes along with a time-relative interpretation. Clearly, in application to abstract objects, *exist* with present tense does not have a time-relative interpretation. The time for the time-relative interpretation may also be conveyed by a temporal adverbial. In the case of a space-relative interpretation of an existence predicate, a spatial adverbial is the only way of indicating a space-relative interpretation.

The spatial or temporal modifier of *occur* simply locates the event described by the verb, an occurrence, at the time or place in question. This is what spatial modifiers generally do when they act as adjuncts of verbs. By contrast, *exist* exhibits a particularly interesting behavior with respect to spatial modifiers, noted and discussed by Fine (2006), which indicates that *exist* with a location modifier

30 Quantified plurals are better with *exist* than (definite) singular NPs:

(i) a. Only ten old Victorian houses exist in this neighborhood.
b. ??? This old Victorian house exists in this neighborhood.

I will have to leave an explanation of this difference to future research.

involves a rather different semantics.[31] One generalization is that *exist* does not allow for spatial modifiers with material objects:

(26) a.??? The building we talked about exists in another city.
b. ??? Notre Dame does not exist in Germany.
c. ??? Every monument we talked about exists in this country.

This means that enduring objects cannot engage in space-relative existence. However, there are some entities that engage in space-relative existence and thus a spatial analogue of endurance. Such objects fulfill the same condition with respect to space as enduring objects fulfil with respect to time. Space-relative existence is well-reflected in English, with existence predicates modified by locative adverbials. Let us look at what sorts of entities allow for *exist* with location modifiers.[32] First, languages and illnesses may engage in space-relative existence:

31 In fact, in the relevant data with *exist*, the location adverbials have the function of complements, not adjuncts, of *exist*. They thus have the same status as *in the house* below, where it is a complement of was:

(i) John was in the house.

Location modifiers that have the status of adjuncts, by contrast, simply act as predicates of the event argument of the verb, as is the case with *occur* as in (25c).
Location modifiers as adjuncts are actually impossible with *exist*, and that is because location modifiers as adjuncts are impossible with stative verbs in general. This is due to a generalization known as the 'Stative Adverb Gap' (Katz (2003), Maienborn (2007), Moltmann (2013b)).
32 For entities permitting space-relative existence, Fine gives the example of a composite aroma of coffee and vanilla whose presence at a location, he argues, requires the presence of both the aroma of vanilla and the aroma of coffee at that location. This example is problematic, though, since aromas do not go along very well with the existence predicate *exist*:

(i) ?? The aroma exists in that room.

The reason why aromas do not go along with *exist* appears to be an ontological one. Aromas as particulars simply cannot be wholly present at different locations and thus cannot have a location-relative existence. Only aromas as kinds can, as in the examples below:

(ii) a. This kind of perfume does not exist in France anymore.
b. This kind of aroma only exists in oriental countries.

Ontologically, aromas as particulars arguably are tropes without a bearer, that is, mere spatio-temporally located features. Tropes in general do not go along very well with space-relative existence predicates:

(27) a. This dialect does not exist in this region anymore.
b. This illness / Syphilis does not exist in Europe anymore.

Moreover, kinds do – more precisely, kinds that are the semantic values of bare (that is, determinerless) plurals and mass nouns when acting as kind terms:[33]

(28) a. Giraffes exist only in Africa.
b. Wild ponies do not exist in Germany.
c. Pure air does not exist in China anymore.

One may think that bare plurals and mass nouns with *exist* are existentially quantified indefinites, reinforcing the contribution of *exist*. However, it is easy to see that bare plurals and mass nouns with *exist* are kind terms and not existentially quantified indefinites. One indication is the behavior of anaphora, as in (29) (which make reference to the entire kind not just some particular individuals); another is the applicability of typical kind predicates like *widespread* as in (30):

(29) a. Dinosaurs do not exist. But they once did exist.
b. Three dinosaurs do not exist. * But they (three dinosaurs or other) once did exist.

(30) Dinosaurs, which used to be widespread in Europe, no longer exist.

It is also significant that *exist* applies to kinds of entities of any sort, including kinds of events:

(31) Political protests do not exist in Bhutan.

(iii) ?? The greenness of the plants exists everywhere in the garden.

Sounds and physical fields for Fine are also entities able to engage in space-relative existence. I find examples with sounds even more problematic than aromas. Sounds as particulars accept neither location-relative nor time-relative existence predicates:

(iv) a. ?? The sound exists throughout the house.
b. ?? The sound we heard five minutes ago still exists.

Sounds as particulars could hardly be present at different spatial locations at once. That sounds do not allow for time-relative existence predicates is no surprise in view of Strawson (1959) point that sounds do not come with criteria for reidentification over time.

33 For the view of kinds acting as semantic values of bare plurals and mass nouns see Carlson (1977) and Chierchia (1998).

This again confirms that bare plurals with *exist* are terms standing for entities that are kinds rather than acting as existential quantifiers.

Why can languages, illnesses, and kinds engage in space-relative existence? What characterizes the entities able to engage in space-relative existence is that they are abstract entities, with an abstract part structure, but yet have concrete manifestations in which all the abstract parts generally are manifested. An illness, for example, may be viewed as a collection of medical conditions which (at least to a great extent) have to be manifested in a patient exhibiting the illness. Languages can be viewed as abstract systems that need to be instantiated in a speaker knowing the language.[34]

How would this apply to kinds? *Exist* can apply to a kind k relative to a spatial location l because k is instantiated at the various (or at least relevant) sublocations of l. This requires a particular conception of kinds on which a kind can be completely present at different locations at once. There is a conception of kinds according to which kinds are just the pluralities of their instances (including perhaps possible instances), a view I argued for in Moltmann (2013a). But that would not permit a kind to be present at several sublocations at once. Rather it requires an abstract conception of a kind, as something like a collection of constitutive features. Only that allows a kind to be completely present at a location l just in case there is a particular individual exhibiting all the constitutive features of the kind at l. A kind then exists at a location l because all its constitutive features are instantiated throughout l.

Complete presence of an entity throughout a location is only possible with abstract entities, entities that have abstract parts all of which can be manifested in a concrete location. Existence in space requires abstract objects able to have concrete manifestations in some sense or other.

Fine's (2006) way of conceiving of existence at a location is hardly applicable to abstract objects of this sort. There is no way, for example, in which the sum of two languages could be at a location just in case one of them is, and so for the sum of two kinds. The restrictions on the sorts of entities that can engage in space-relative existence instead support the more traditional complete presence conditions, whatever the difficulties may be of making that condition entirely precise.

34 Of course, a speaker may only partially know the language, but parts of a language are entities strongly dependent on the whole of the structure of the language, and thus partial knowledge of a language may be viewed as an instantiation of at least the structural whole of the language.

6 The existence predicate *obtain* and the category of condition-like objects

Exist and event predicates like *happen* are not the only existence predicates in English. Another existence predicate is *obtain* (and the near-synonym *hold*). *Obtain* is an existence predicate that is generally interchangeable with *exist*, but it can apply only to what I call 'condition-like objects'. As in the case of *exist*, condition-like objects permit both space- and time-related uses of *obtain*. The reason they permit space-related uses is again, I will argue, that they have abstract parts, though of a different sort than kinds, languages, and illnesses.

Obtain applies to two kinds of condition-like objects. First, it applies to entities we refer to as 'states', 'situations', or 'conditions':[35]

(32) The state / situation / condition still obtains.

Second, it applies to entities such as permissions, obligations, and laws, the 'modal products' of directive or declarative speech acts:[36,37,38]

(33) a. The permission / obligation still obtains.
b. A new law now obtains.

35 *Obtain* also applies to facts (though time-relative and space-relative *obtain* does not apply to facts since the canonical description of facts about concrete entities (Section 4) includes a (explicit or implicit) location specification):

(i) The fact that S obtains.

36 For the notion of a (modal) product of an illocutionary act, a notion that derives from Twardowski, see Moltmann (2017b).

37 *Exist* is actually not that good with some of the modal products, for example permissions :

(i) The permission to skip the meeting still obtains / ??? exists.

This means that *exist* is subject to another restriction, though it is not obvious what that would be.

38 Not only *obtain* and *exist* are applicable to normative condition-like entities such as laws, but also *is valid*:

(i) The law still obtains / exists / is valid in some countries.

Is valid, as was mentioned in Section 2.1., can act as an existence predicate, but as such it is restricted to normative products established by declaration (laws, rules, offers etc).

Obtain does not apply to material objects, persons, or events, and it does not apply to abstract objects of the sort of mathematical objects, properties, or propositions. Moreover, it fails to apply to kinds:

(34) a. ??? The house / The person / The smallest prime number / The event obtains.
b. ??? Giraffes obtain.

Like *exist*, *obtain* has a time-relative use, as in (32, 33) as well as a space-relative use, as below:

(35) a. The state / situation / condition / law still obtains in various regions of the country.
b. The state of emergency / The same situation / The same condition now obtains in Arizona.

It is remarkable that *exist* applies to states, when it could not apply to events. It is indicative of a fundamental difference between events and states, namely that events have temporal parts, but states don't.

What characterizes the entities to which *obtain* is restricted is that they are constituted by particular conditions holding, that is, by specific properties or relations holding of an object or a number of objects, possibly at a time and a spatial location. That is, they are entities whose existence and identity depends on those particular conditions being in place. This is why I call them *condition-like entities* and the conditions constitutive of them *constitutive conditions*.[39]

Some condition-like entities go along with canonical descriptions, that is, descriptions that display exhaustively the nature of the entities they stand for in terms of such conditions. Facts have canonical descriptions of the form *the fact that S*, states have canonical descriptions of the form *the state of NP*s being *VP*, and conditions have canonical descriptions of the sort *the condition of NP*'s *being VP*. The canonical descriptions make explicit the properties or relations and the relevant objects and possibly locations that are constitutive of the condition-like entity.

The condition-like entities that come with canonical descriptions are entities that fall under Kim's account of 'events'(Kim (1976)) , which is in fact an account of (non-worldly) facts (Moltmann (2013a,b)). The Kimean account consists in a specification of the existence and identity conditions of events, or rather facts, on

39 The way *obtain* differs from *exist* may be attributed to the particular notion of 'presence' it involves: *exist* requires presence in the sense of spatial or temporal locatedness, whereas *obtain* requires the more specific notion of presence in the sense of a property being true of an object relative to a location.

the basis of a property, an individual, and a time, as below, where properties are taken to be functions from times to sets of entities:

(36) The Kimean account of events (facts)
(i) For a property P, an object o, and a time t, $f(P, o, t)$ obtains iff $o \in P^t$.
(ii) For properties P, P', o, o', $f(P, o) = f(P', o')$ iff $P = P'$, $o = o'$.

For states, condition (36ii) should be replaced by the one below, with location-relative *obtain* relativized either to a temporal or spatial location:[40]

(37) For a property P and an object o, $f(P, o)$ obtains at a location l iff $o \in P^l$.

(36i) and (37) together of course imply (38):

(38) For a property P, an object o, and a time t, $f(P, o)$ obtains at t iff $f(P, o, t)$ obtains.

The Kimean account amounts to an implicit definition of entities. It introduces entities that have just those properties as intrinsic properties that are specified by the account itself (Moltmann (2013b)). This is why it defines facts and states as nonwordly entities. Thus, given (36, 37), facts will have neither a temporal nor a spatial location, and states will have only a temporal location. In addition, given (36, 37), facts and states need not be based on a specific property, but may be based on a non-natural, or disjunctive one.

Location-relative *obtain* involves the same application condition as location-relative *exist*, namely a condition of 'complete presence' at the relevant sublocations of the relevant location. What exactly does the complete presence of a situation, state, or condition at a time or location consist in, that is, what would count as the parts of a condition-like entity that would have to be present at the relevant sublocations? The objects and times from which condition-like entities are obtained (in the 'Kimean' way) certainly do not count as parts of such entities. This is reflected in the fact that they are not treated as parts by part-related expressions of natural language: *part of the situation*, *part of the condition*, or *part of the state* can never 'mean' a participant or location of the situation, condition, or state. Moreover, condition-like entities, unlike material objects, do not have spatial parts, and unlike events, they do not have temporal parts. Rather their parts are their constitutive subconditions and it is with them that the complete presence at sublocations needs to be fulfilled when *obtain* applies relative to a temporal or

40 This gives a characterization of states as 'Kimian' states (Maienborn (2007)) or 'abstract' states, as I call them in Moltmann (2013b), rather than 'Davidsonian' states (Maienborn (2007)) or 'concrete' states (Moltmann (2013b)).

a spatial location. Thus, for a situation, condition, or state *e* to obtain relative to a location *l* means that all the constitutive subconditions of *e* are fulfilled at all the relevant sublocations of *l*. Because condition-like entities have neither temporal nor spatial parts, they can be completely present at different times as well as different places.

Modal products do not exist in virtue of an implicit definition, but rather in virtue of the declared validity of the relevant normative condition.[41] Again what counts as parts of a normative product is its constitutive subconditions, not temporal or spatial parts. This explains why modal products allow for both time-relative and space-relative applications of *obtain*.

What is common to both kinds of condition-like entities is that their existence at a time or space depends on those conditions being in place at that time or that space. For a condition-like entity to obtain at a location, either all the various things need to happen at the location in virtue of which the constitutive condition (with all its subconditions) holds or else the constitutive condition needs to have been put in place by declaration for that location. Either way, the condition-like entity will need to enjoy complete presence throughout the location as long as the constitutive condition, with all its subconditions, holds.

For condition-like entities, merely intentional counterparts are much easier to accept than they are for material objects. Their intentional counterparts are simply their constitutive conditions (possibly together with the relevant time or location). For condition-like objects, the treatment of existence predicates as predicates true of existent and false of nonexistent objects thus has a particularly plausible basis.

7 Location-independent uses of existence predicates and summary

In its location-relative use, *exist* has a nontrivial meaning, describing a state of an entity *d* that involves the complete presence of *d* at the sublocations of the temporal or spatial location in question. *Occur*, by contrast, tracks the temporal loca-

41 States are actually condition-like entities that may be of either kind, as states based on empirical facts (about the time or spatial location, or the world) and as states based on normative conditions or conditions resulting from 'declarations' (which may or may not be restricted to a time or a spatial location). The state of someone's mind or health is a state of the first kind, as are habits; a state of war, a requirement and a law are condition-like entities of the second kind. The first kind of state holds in virtue of what is taking place at the relevant location; the second kind of state holds by declaration or whatever may ground normative conditions.

tions of subevents of an event *e*, thereby describing another event (an occurrence) that reflects the mere temporal structure of *e*. But what is the meaning of location-independent uses of existence predicates? In particular, what is the meaning of present-tense *exist* that would not involve a time-relative interpretation?

It is important to note that location-independent uses of existence predicates still impose the same restrictions on the sorts of entities they can apply to. In particular, time-independent *exist*, like time-relative *exist*, is inapplicable to events. The preservation of the type restrictions indicates that the time-independent meaning of *exist* is derivative upon its time-relative meaning and motivates a semantics of time-independent *exist* involving universal quantification over all times:

(39) For an entity d, exist(d) = 1 iff exist(d, t) for all times t.

Exist with that meaning could not apply to events: it would require the complete presence of an event at all times, which is impossible. But *exist* with that meaning can apply to abstract objects since abstract object can naturally be considered completely present at all times. Time-independent *exist* thus can be derived from time-dependent *exist* and *occur*. But can space-independent exist be derived from space-relative *exist*? This would require complete presence everywhere of entities to which space-independent *exist* applies. However, this is impossible for material objects to fulfill. This means is that space-relative *exist* does not have a space-independent correlate, unlike time-relative *exist*, or in other words, location-independent *exist* involves only universal quantification over times, not spatial locations, and thus is based on time-relative *exist* only, as in (39).[42]

Let me then summarize the selectional restrictions of existence predicates on their location-independent and location-dependent uses as follows:

(40) Summary of the selectional restrictions of existence predicates

a. *Exist*
Location-independent use:
Material objects (including artifacts and organisms), states, conditions, laws
Location-dependent use:
Time-dependent: same as location-independent use
Space-dependent: kinds (including kinds of events), illnesses, languages

b. *Occur* (*happen, take place*)
Location-dependent and location-independent uses: events (of certain sorts

42 This is different for *obtain*: a fact can satisfy a vacuous relativization to a time just as it can satisfy a vacuous relativization to a space.

c. *Obtain*
Time-independent: facts, states, conditions, laws
Location-dependent: states, conditions, laws

These selectional restrictions, we have seen, follow from the meanings of *exist*, *occur*, and *obtain* with their possible relativization to a time or space. *Obtain* imposes the more specific condition of the relevant constitutive condition holding at the location *l*, rather than the general condition of being present at a location *l*.

8 The plausibility of intentional objects and nonexistence as the failed intentionality of intentional objects

Existence predicates were characterized as a class of predicates with a particular semantic property, their acceptance of apparently empty subject terms. The paper has argued that such terms in fact stand for intentional (nonexistent) objects. The variety of existence predicates, applying to different sorts of entities, gives additional plausibility to that view. Certain types of objects come with a very intuitive intentional counterpart. In particular, a range of entities to which *obtain* applies have easily acceptable 'non-obtaining', intentional counterparts. Laws, for example, appear to have proposition-like entities as intentional counterparts. For a law, existence amounts to a declaration of a proposition. States, situations, or states of affairs, moreover, are entities to which 'obtaining' clearly adds something: it requires the constitutive condition of the state, situation, or state of affairs to be true of the object in question (at the relevant location). A distinction between existing and merely intentional entities is quite intuitive also for events. The German counterpart of *take place*, *stattfinden*, for example, requires the event term to describe an event that has been planned before and thus had a previous intentional counterpart; it cannot just apply to an event term that fails to refer to an event or with which an agent had failed to refer to an event (*Der Tod des Mannes fand nicht statt* 'The death of the man did not take place'). Finally, location-relative *exist* when it applies to kinds involves the attributes making up the essence of the kind for the fulfilment of the complete presence condition, which then amounts to complete instantiation. *Exist* on that use thus presupposes a distinction between the essence of a kind (which does not require instantiation) and an 'existing', that is, instantiated kind. Thus, looking at a greater range of entities and existence

predicates, there appears to be a significant plausibility for there being entities with different modes of being than existence.[43]

McGinn (2000) gives a somewhat different account of negative existentials, based on the view that apparently empty terms stand for intentional objects. For McGinn, intentional objects are entities constituted by failed intentionality. Only when a speaker when using a term fails to refer to an actual object will the term stand for an intentional object, that is, an object constituted by the unsuccessful act of reference. *Exist*, for McGinn, is necessarily false of intentional objects, and that is because of the particular nature of such objects as merely intentional objects. Let me call such a potential use of *exist* the *reference-related use* of *exist*. Positing intentional entities as entities that unlike any other necessarily fail to exist may seem like a highly problematic view (Van Inwagen (2008)). But there may be a way of making the view somewhat more plausible, and that is, by taking *does not exist* when applied to an intentional object to have a derivative meaning, specifying the failed intentionality constitutive of such an object. This means that 'nonexistence' would be an essential property of intentional objects only derivatively, derivative upon a nonessential property (failed intentionality) of the intentional act constitutive of the intentional object.

If *exist* can be understood that way, a number of predictions are made. First, *exist* should not be applicable relative to a time, but only time-independently. Furthermore, there would be no reason for *exist* to impose selectional restrictions on the types of entities it applies to (since the type restrictions were only to guarantee the applicability of the location-relative use). It is not clear that intuitions support a reference-related use of *exist*. The following examples should be acceptable, but certainly not all speakers agree that they are:

(41) a. Does the medieval war described in the book exist?
b. The Third World War does not exist.

With a use of *does not exist* specifying failed intentionality, one would expect another time-relative use, relating to the time of the intentional act rather than the intended time of existence of the object the agent tries to refer to. On such a use,

43 The close connection between instantiation and existence might make it tempting to reconsider an account of existence statements with singular terms, namely, according to which *exist* expresses instantiation, relating an individual concept to a material manifestation at a location. Of course, the individual concept would not be given by a predicative expression, but be associated with the use of a singular term in the particular context of an existence statements. I will not pursue this option further, though.

(42a, b) should be acceptable, stating, basically, that yesterday's reference act was successful:

(42) a. Yesterday Socrates did exist.
b. Yesterday the ancient philosopher Bill mentioned did exist.

But here the intuition is strong that a reading is unavailable on which the sentence could be true (when uttered by a contemporary speaker). The reference-related use of *exist* thus is certainly not the main use of *exist*.[44] Rather it is the location-related use, and the location-independent use is derived from that.[45]

9 Conclusion

This paper has given support for several generalizations about the expression of existence in natural language. First, quantifiers in natural language are neutral regarding existence and nonexistence (though they can also be used in a non-neutral way). They can be used for merely intentional (nonexistent) objects equally well as for actual objects.

The paper gave specific further motivations for merely intentional objects. There are constructions that require merely intentional objects for their compositional semantics. Moreover, while merely intentional entities are generally considered problematic, for condition-like entities the distinction between actual and merely intentional entities is much more intuitive and acceptable than in the case of material objects, which the philosophical discussion has generally limited itself to.

Natural language reflects different ways of being, but not with different quantifiers, but rather with different existence predicates, a semantically distinguished class of predicates. At least in English (and related European languages) existence predicates express different ways for entities to relate to time and space, rather than notions such as fundamentality, ontological dependence or derivativeness, or mind-dependence. The predicate *exist* itself English display a close connection

44 The nominalization *existence*, though, appears have a reference-related use, as in *the existence of the event described in the book*. Alternatively, *existence* may just convey the reflective notion of existence, see Fn 12.

45 Note that a reference-related use is entirely unavailable for other existence predicates than *exist*.

between existence and endurance, the latter is just what time-relative *exist* expresses.

Location-relative *exist* appears to reflect the traditional notion of 'complete presence' throughout a location. What is important about this notion is that it is applicable not just to material objects, but also to abstract objects that have concrete manifestations such as languages, illnesses, and kinds, as well as to condition-like entities such as states, laws, and conditions, the sorts of entities that can engage in space-relative existence.

The notion of existence reflected in location-relative *exist* involves the recurrence of the essential parts or features of an entity, the preservation of its identity, across locations. This notion is inapplicable to entities of the sort of events and tropes, which in that sense then have a lesser degree of being. In that way, one may say, existence predicates in English distinguish between degrees of being after all.

It has sometimes been argued that linguistic intuitions about the verb *exist* should not be taken too seriously, for making either a semantic or a philosophical point, since *exist* is a relatively recent verb and tied to a more 'technical' use in philosophical contexts. This caution turns out to be entirely in error. The paper has shown that *exist* displays a surprising and systematic semantic behavior and as such does not in fact convey the reflective notion of existence. Instead, *exist* conveys a notion of endurance or its space-relative analogue, or else a location-independent notion that is derivative upon the location-dependent notion. These are the very same notions that the existence predicate *obtain* conveys, except that the latter involves a more restricted notion of presence at a location. The existence predicate *obtain* displays the very same semantic behavior, but is restricted to condition-like entities. The linguistic intuitions associated with *exist* thus display a more general concept of location-relative existence as part of the metaphysics implicit in language, rather than peculiar features of a somewhat special lexical item.

Bibliography

Austin, J. L. (1979), "Unfair to Facts", in *Philosophical Papers*, New York: Oxford University Press.

Azzouni, J. (2010), *Talking about Nothing: Numbers, Hallucinations, and Fictions*, Oxford: Oxford University Press.

Bach, E. (1986), "Natural Language Metaphysics", in *Logic, Methodology, and Philosophy of Science VI*, edited by B. Marcus et al., Amsterdam et al.: North Holland, 573–595.

Carlson, G. (1977), "A Unified Analysis of the English Bare Plural", in *Linguistics and Philosophy*, 413–457.
Casati, R. and Varzi, A. (2005), "Events", in *Stanford Encyclopedia of Philosophy*.
Chierchia, G. (1998), "Reference to Kinds across Languages", in *Natural Language Semantics*: 6, 339–405.
Cresswell, M. J. (1986), "Why Object Exists, but Events Occur", in *Studia Logica*: 45, 371–375.
Dretske, F. (1967), "Why Events Cannot Move", in *Mind*: 76, 479–492.
Fine, K. (2006), "In Defense of Three-Dimensionalism", in *Journal of Philosophy*: 103(12), 699–714.
Fine, K. (2017), "Naïve Metaphysics", in *Philosophical Issues*: 27(1), edited by J. Schaefer, 98–113.
Geach, P. T. (1968), "What Actually Exists", in *Proceedings of the Aristotelian Society*, Supp.: 42, 7–16.
Hacker, P. M. S. (1982a), "Events, Ontology, and Grammar", in *Philosophy*: 57, 477–486.
Hacker, P. M. S. (1982b), "Events in Time and Space", in *Mind*: 91, 1–19.
Hawley, K. (2001), *How Things Persist*, Oxford: Oxford University Press.
Heller, M. (1990), *The Ontology of Physical Objects*, Cambridge: Cambridge University Press.
Katz, G. (2003), "Event Arguments, Adverb Selection, and the Stative Adverb Gap", in *Modifying Adjuncts*, edited by E. Lang, Berlin: Walter de Gruyter.
Kim, J. (1976), "Events as Property Exemplifications", in *Action Theory*, edited by M. Brand and D. Walton, Dordrecht: Reidel.
Lewis, D. (1986), *On the Plurality of Worlds*, Oxford: Blackwell.
Maienborn, C. (2007), "On Davidsonian and Kimian States", in *Existence*: *Semantics and Syntax*, edited by I. Comorovski and K. von Heusinger. Springer, 107–130.
McDaniel, K. (2009), "Ways of Being", in *Metametaphysics*: *New Essays on the Foundations of Ontology*, edited by D. J. Chalmers, D. Manley and R. Wasserman . Oxford: Oxford University Press.
McDaniel, K. (2010a), "A Return to the Analogy of Being", in *Philosophy and Phenomenological Research*: 81(3), 688–717.
McDaniel, K. (2010b), "Being and Almost Nothingness", in *Nous*: 44(4), 628–649.
McDaniel, K. (2013), "Degrees of Being", in *Philosophers' Imprint*: 13(19), 1–18.
McGinn, C. (2000), *Logical Properties*, Oxford: Oxford University Press.
Merricks, T. (1995), "On the Incompatibility of Enduring and Perduring Objects", in *Mind*: 104, 523–531.
Miller, B. (1975), "In Defense of the Predicate Exist", in *Mind*: 84, 338–354.
Miller, B. (1986), "Exists and Existence", in *The Review of Metaphysics*: 40, 237–270.
Miller, B. (2002), "Existence", in *Stanford Encyclopedia of Philosophy*.
Moltmann, F. (2013a), *Abstract Objects and the Semantics of Natural Language*, Oxford: Oxford University Press.
Moltmann, F. (2013b), "On the Distinction between Abstract States, Concrete States, and Tropes", in *Genericity*, edited by A. Mari, C. Beyssade and F. Del Prete, Oxford: Oxford University Press, 292–311.
Moltmann, F. (2015), "Quantification with Intentional Verbs and with Intensional Verbs", in *Quantifiers, Quantifiers, Quantifiers*, edited by A. Torza, Dordrecht: Synthese Library, Springer, 141–168.
Moltmann, F. (2017a), "Natural Language Ontology", in *Oxford Encyclopedia of Linguistics*.

Moltmann, F. (2017b), "Cognitive Products and the Semantics and Attitude Verbs and Deontic Modals", in *Act-Based Conceptions of Propositional Content*, edited by F. Moltmann and M. Texto, New York: Oxford University Press.
Muyskens, R. (1989), *Meaning and Partiality*, Chicago: Chicago University Press.
Parsons, T. (1980), *Nonexistent Objects*, New Haven: Yale University Press.
Priest, G. (2005), *Towards Nonbeing*, Oxford: Oxford University Press.
Russell, B. (1905), "On Denoting", in *Logic and Knowledge*, edited by R. C. Marsh, London: Allen & Unwin.
Sainsbury, R. M. (2005), *Reference without Referents*, Oxford: Oxford University Press.
Salmon, N. (1987), "Existence", in *Philosophical Perspectives*: 1, 49–108.
Salmon, N. (1998), "Nonexistence", in *Nous*: 32(3), 277–319.
Sider, T. (2001), *Four-Dimensionalism*, Oxford: Clarendon Press.
Simons, P. (1994), "Particulars in Particular Clothing. Three Trope Theories of Substance", in *Philosophy and Phenomenological Research*: 54(3), 553–575.
Strawson, P. (1950), "Truth", in *Proceedings of the Aristotelian Society*: 24, 129–156. Reprinted in Strawson, P. (1971), *Logico-Linguistic Papers*, London: Methuen.
Strawson, P. (1959), *Individuals. An Essay in Descriptive Metaphysics*, London: Methuen.
Turner, J. (2010), "Ontological Pluralism", in *Journal of Philosophy*: 107(1), 5–34.
Van Inwagen, P. (1998), "Meta-ontology", in *Erkenntnis*: 48, 233–250.
Van Inwagen, P. (2008), "McGinn on Existence", in *Philosophical Quarterly*: 58(230), 36–58.
Wiggins, D. (1980), *Sameness and Substance*, Oxford: Blackwell.
Williams, D. C. (1953), "On the Elements of Being", in *Review of Metaphysics*: 7, 3–18.
Zalta, E. N. (1983), *Abstract Objects: An Introduction to Axiomatic Metaphysics*, Dordrecht: Reidel.
Zalta, E. N. (1988), *Intensional Logic and the Metaphysics of Intenstonality*, Cambridge, MA: MIT Press.

Kevin Mulligan

Modes of Being and the Mind

Abstract: Are there modes or kinds of being as opposed to modes or kinds of beings? Many heirs of Brentano not only endorsed modes of being but happily distinguished a large number thereof. Those who resisted the very idea of modes of being, such as the early Husserl and Marty, relied on versions of the view that being or existence is not a first-order property. According to many of their twentieth century fans, modes of being have one thing in common: each mode of being is correlated with a type of mental act or state. It is argued in this paper that even if modes of being are rejected the properties and relations which were wrongly classified as modes of being may well be correlated with different types of mental acts and states.

1 Introduction

A striking difference between Continental Philosophy (CP) and Analytic Philosophy (AP), as Peter van Inwagen was the first to point out, is the popularity of modes (kinds, ways) of being in CP and their almost complete absence in AP, where versions of Frege's (and then Russell's) account of existence or being as a second-order property have been so popular. Geach and Kenny, it is true, had pleaded in favour of Aquinas' actuality or *esse naturale*, and *esse intelligibile*. And Vallicella and, more recently, McDaniel have vigorously defended the cause of modes of being. Their popularity and multiplication in CP, it is often thought, goes back to Heidegger. In fact, it is Scheler who is at the beginning of this development within early phenomenology and, as we shall see, it is a compatriot of van Inwagen, Vallicella and McDaniel, namely, William James, who seems to be one of the main precursors of Scheler and Heidegger. The enthusiastic multiplication of of being has been a feature of CP (Sartre, Gadamer, Foucault) ever since.

But the multiplication was preceded by defences of a mere handful of modes of being by such heirs of Brentano as Husserl and Meinong. Even earlier, some of Brentano's heirs followed Brentano in clearly rejecting the very idea of modes of being altogether. This is true of Husserl in 1894 and Marty in 1908. It is almost certainly no accident that these two rejections of modes of being went hand in hand with endorsements and defences of alternatives to the view that being or existence is a first-order property. But Husserl and Marty do not have the same account of what it is to be. Husserl's account is taken over from Bolzano, who understands

https://doi.org/10.1515/9783110664812-010

> There are A's

as

> The idea (in itself) A has objectuality (*Gegenständlichkeit*) (Bolzano (1929), II §137)

Simplifying slightly, we may say that, according to Bolzano, for there to be A's is for the concept A to have objectuality. What Bolzano calls a concept is not what Frege, rather strangely, was later to call a concept and employ in his account of existence. Bolzano's concepts play rôles like those played by what Frege calls sense. In 1894, in a manuscript in which he endorses a version of Bolzano's account, Husserl asserts: "There are not different modes of existence and validity" (Husserl (1979), p. 326). Marty's rejection of modes of being, on the other hand, is motivated by his endorsement of Brentano's account of existence:

> we say an object exists, if the accepting directed towards it is immediately or mediately evident (Brentano (1978), pp. 146–147)

by which he really means something like this: an object exists if the accepting-directed-towards-it is knowable as evidently correct, immediately or mediately. Marty's forceful 1908 rejection of modes of being for what is real and for what is non-real argues that they are *superfluous* if one really understands what it is to be:

> there is a difference in the what (*Was*) between the two, that seems to me to be enough. Of course, if it were a fact that there are different ways of being, this objection would not apply. But is it a fact? I cannot find that this is the case … "Being" means … what deserves to be accepted, nonbeing what deserves to be rejected.[1]

Marty's campaign against modes of being is, as far as I can see, comparable (in tone and in other respects) only to that mounted much later by van Inwagen.[2] It has two interesting features. First, Marty tracks in detail the mistakes bound up with the endorsement of something like the mode of being often called intentional

1 Marty, 1976 (1908), pp. 323–324. Marty goes on to make a curious concession to friends of modes of being. He says that "in a certain sense" there are three ways of being, namely being necessary, being possible and being identical with something (Marty, 1976 (1908), p. 324). The ways of being he rejects are intentional being, being real and a particular form of being non-real, obtaining (*bestehen*, *Bestand*). On the concession, cf. Taieb (forthcoming).

2 Van Inwagen (1981, 2014) (especially ch. 3, "Being, existence, and ontological commitment"); cf. "Need I add that I have not the slightest notion of what a *Seinsart* could possibly be? The one thing I am sure of is that it is something even more absurd than a level" (Bergmann, 2004 (1967), p. 418, cf. 3), Specht (1967), Grossmann (1984).

or intelligible being, in discussions of numerous historical figures, an endorsement of which he himself, he confesses, was once guilty. Second, in his critique of the very idea of intentional being and elsewhere he introduces a type of objection which was to have almost as great an effect on twentieth century philosophy as the propaganda in favour of modes of being by Scheler and Heidegger. To take seriously mental being, Marty argues, is to be misled by a fictitious picture (*Bild*), that of a *Doppelgänger*; it is to be tempted by language and to succumb. Many of the misleading pictures, analogies and fictions fingered by Marty figure prominently in the writings of the later Wittgenstein.[3]

In what follows, I first sketch the multiplication of modes of being in the writings of Brentano's heirs (§2) and distinguish between endorsement of modes of being as such and the use of the language of modes of being to refer to regions, domains, provinces and spheres of being. I then (§3) ask what it is, in this tradition, that modes of being are supposed to have in common, other than the supposed fact that they are modes *of* being. The answer, as we shall see, is that, with a few exceptions, the different modes of being are supposed to be *correlated* with different types of mental acts. The interesting thing about this claim is that even if, as I believe, there are no such things as modes of being and what have been called modes of being belong to much more familiar categories, such as that of property and relation, the claim that there are correlations between these properties and types of mental act retains great plausibility. Finally (§4), I consider the accounts by Brentano's heirs of the relation between essence and modes of being and put forward a second way of distinguishing between the properties wrongly identified as modes of being and other properties.

2 The Multiplication of modes of being

Many ideas which were to play an important role in phenomenology are sketched out in James' *The Principles of Psychology* and elsewhere. These include ideas about ontology. James is a great fan of types of existence. He not only anticipates the multiplication of modes of being in CP, he also, as we shall see (§3), anticipates a way of arguing for or identifying modes of being which was to be very important in CP. In Chapter 21 of his *Psychology*, "The Perception of Reality", James, perhaps influenced by Lotze, says:

3 Cf. Mulligan (2019).

> A dream-candle has existence, true enough; but not the same existence (existence for itself, namely, or *extra mentem meam*) which the candles of waking perception have. A dream-horse has wings; but then neither horse nor wings are the same with any horses or wings known to memory.
> Habitually and practically we do not *count* these disregarded things as existents at all. For them *Vae victis* is the law in the popular philosophy; they are not even treated as appearances; they are treated as if they were mere waste, equivalent to nothing at all. To the genuinely philosophic mind, however, they still have existence, though not the same existence, as the real things. As objects of fancy, as errors, as occupants of dreamland, etc., they are in their way as indefeasible parts of life, as undeniable features of the Universe, as the realities are in their way. The total world of which the philosophers must take account is thus composed of the realities *plus* the fancies and illusions.

By 1910, in the second edition of *Über Annahmen*, Meinong is referring to "the essential difference between" two "kinds of being" (*Seinsarten*):

> What exists (*existiert*) must exist at a certain time; existence is attached to time. So little is that the case with obtaining (*Bestand*) that the sort of thing that by nature can only obtain (*bestehen*) and cannot exist – something ideal, in other words – tolerates no time determinations whatsoever. To attribute a time to the difference between red and green would have scarcely any more sense than to call a tone white or black … [It is a] remarkable fact that in the domain of the obtaining there is no causal investigation, although such investigation is inseparable from existence (Meinong (1983), p. 59 – tr. modified, Meinong (1977), p. 75)

Meinong here claims that kinds of being and the modal properties of these and of what enjoys these kinds of being are determined by natures: it is the nature of the difference between red and green which is responsible for the fact that it must enjoy the mode of being of obtaining and cannot enjoy the mode of being of existence. Modes of being and natures or essences are not, then, the same sort of thing and the latter are prior to the former. Both claims are endorsed by many phenomenologists but not, as we shall see, by Heidegger (§4).

Meinong even wonders whether there is something like a third kind of being:

> In this sense, "there are (*es gibt*)" also objects which are (*sind*) not [do not exist or obtain], and I have called this the "*Aussersein* of the pure object", a somewhat barbaric word-formation but one which it is hard to improve on. – The term arose in connection with the effort to interpret that peculiar "there are" that it seems impossible to detach from the objects which are furthest from being without having recourse to a new, third kind of being in addition to existence and obtaining. … I must mention the possibility that there is a third something beyond existing and obtaining to which no one extends the term being but which

> would have to be characterised as something being-like (*Seinsartiges*) in a particularly wide sense of the word.[4]

In 1913, Husserl, too, is happily referring to modes of being (*Seinsarten*), to "immanent or absolute being", the being of consciousness, and to "real", "transcendent" being (Husserl (1950) §49). Marty clearly thinks that already in 1900-1901 Husserl accepts modes of being. He also takes Meinong to be so committed in 1902. Even Bolzano, Marty thinks, accepts modes of being. In the case of Husserl, it is true that already in his *Logical Investigations* he refers to what he will later call modes of being, to "the difference between ideal being (*idealem Sein*) and real being (*realem Sein*) (Husserl (1984) II, §8).

The small number of modes of being allowed for by Meinong and Husserl was rapidly augmented by Scheler and later phenomenologists:

List 1:

– Being *wirklich* (actual), *individuell*, real (*real*)	Husserl, Ingarden, Scheler, Heidegger
– Being ideal, subsisting	Husserl, Meinong, Ingarden, Heidegger
– Endurance (*verharren*)	Ingarden
– the mode of being of momentary events	Ingarden
– the mode of being of processes	Ingarden
– Being alive	Scheler, Heidegger, Conrad-Martius
– Being *vorhanden*	Heidegger
– Being *zuhanden*	Heidegger
– *Existenz*	Heidegger
– Being valuable	Scheler, Hartmann, Ingarden, Sartre
– Being psychological	Husserl, Scheler
– Being-*pour-soi*	Sartre
– Being physical	Husserl, Scheler
– Being-*en-soi*	Sartre
– Being a person	Scheler
– Being a force	Ingarden
– Being possible	Scheler, Spielberg, Ingarden
– Being necessary/probable	Scheler, Spiegelberg
– Being present, past or future	Scheler, Heidegger

4 Meinong (1977), p. 62, and cf. Meinong (1983), pp. 79–80. On Meinong's modes of being, cf. Richard (2017). Already in 1904, Meinong's student, Ameseder, refers to existence and obtaining as modes of being (Ameseder (1904), pp. 78, 79 and 83), mentions the possibility of a "third kind of being" and clearly distinguishes the claim that existence is a way of being which is "qualitatively different" from obtaining and the claim that existence is the being of objects "of a completely different kind" from ideal objects (Ameseder (1904), p. 81).

The list is long but by no means complete. Some of the envisaged modes of being are said to be enjoyed by objects, some by states of affairs, some, for example, being valuable and being past, by both. But Ingarden, for example, seems to think of being present, past and future as modifications (*Abwandlungen*) of being real. Hartmann distinguishes between what have here been called modes of being (*Seinsweisen*) and the *Seinsmodi* of possibility, actuality and necessity. Some items on the list are more clearly formal or topic-neutral (the semantic values of the temporal, modal and value connectives) than others (psychological being, physical being, being alive). Some seem to stand to each other in a relation which is a counterpart of the determinable-determinate relation between properties or qualities. To be real is to endure, to enjoy the mode of being of momentary events or that of processes. To be real is to be alive or psychological or physical.[5]

Many of the friends of modes of being mentioned in the above list seem to be either ignorant of philosophies of being which do not countenance modes of being or do not give explicit accounts of the relation between being, on the one hand, and modes *of* being, on the other hand. This is, as far as I can see, true of Scheler (although he is aware of the need for an account of the relation between what makes true what he calls "existential propositions" and modes of bring), Heidegger in *Sein und Zeit*, Ingarden and Sartre. It is enough to consider the case of Ingarden, the author of by far the most interesting account of modes of being, who remarks *en passant* that "Being or a mode of being … is not a property of a property or of anything else" (Ingarden (1965), §42, p. 96) and asserts that

> ´Existence' (*Existenz*) in general is merely a general idea, the instantiations (*Vereinzelungsbesonderheiten*) of which are the individual modes of being (*Seinsweisen*).[6]

Husserl, we have seen, knew and accepted Bolzano's account of (what Husserl calls) existential propositions:

> Where one judges about existence… "meanings" come in as the judged objects. … This was moreover correctly seen by Bolzano (Husserl (2009), p. 160)

Early and late, he accepts something like Bolzano's account of existence. Why, then, did he give up his early opposition to modes of being and endorse both these modes and Bolzano's account of existence? The question is touched on in 1913:

---

5 On overlap of modes of being, cf. McDaniel (2017).

6 Ingarden (1964), p. 78. It is not clear what, according to Ingarden, the relation is between this claim and the following one: "After the introduction of different modes of being (*Seinsweisen*) … the little word ´is'… in its existential meaning is ambiguous" (Ingarden (1964), p. 67).

> We thus see that consciousness (experience) and real being are in no sense coordinate modes of being (*Seinsarten*). … Both immanent or absolute being and transcendent being are indeed both "seiend" and "object" and each has its objective determining content: but it is evident that what then on either side goes by the name of object and objective determination of an object bears the same name only when we speak in terms of *the empty*, *logical categories*. Between consciousness and reality there yawns an abyss of sense … (Husserl (1950), §49; emphasis mine – KM)

The claim that the property of having objectuality is logical or formal is very plausible, as is the view that having zero extension is a topic-neutral matter. Instead of "formal" or "logical" one might here employ the term van Inwagen uses: "thin". Unfortunately, neither here nor elsewhere does Husserl tell us what the relation is between thin, logical categories and such thick or, to use Ameseder's term, qualitative notions as mental and physical being.

Simplifying only slightly, we have so far distinguished three philosophies of modes of being within the Brentanian tradition. There is the rejection of the very idea combined with, if not motivated by, rejection of the view that existence is a first-order property and endorsement of an alternative, "thin" account of being or existence (early Husserl) or of an account in terms of what it is correct to accept (Marty, 1976 (1908)). There is, secondly, endorsement of modes of being combined with silence about (and perhaps ignorance of) accounts of being like those given by Bolzano, Frege or Brentano and Marty. (Scheler, Ingarden, Heidegger in 1927, Sartre). There is, thirdly, it seems, explicit endorsement of both thin and thick accounts of being, of being as a second-order property and of modes of being (Husserl sometime after 1894). But is this really Husserl's view? Before we try to answer this question, it will be useful to consider a fourth conception of modes of being.

It is to be found, strangely enough, in Carnap's 1928 *Aufbau*. Carnap there introduces his distinctions between different modes of being with the following remark:

> Following an occasionally employed use of language, one may distinguish the different "modes of being" (*Seinsarten*) of the objects of different spheres. This expression makes particularly clear just how completely separated and incomparable objects are which do not belong to the same sphere (*sphärenfremde*) (Carnap (1974), §42).

Carnap in 1928 was familiar with the writings of Husserl and Meinong and presumably with their use of the term *Seinsart*. Presumably he had also come across Husserl's distinctions between different *spheres* and *regions* of being. When Carnap goes on to deploy the notion of modes of being, he turns out to be a multiplier of modes of being, just as much of a multiplier as James, Scheler and Heidegger,

although the way he multiplies modes of being is very different from theirs and his *list* is not theirs

> Classes are constructed from things. These classes do not consist of the things. They *are no beings* (kein *Seiendes*) *in the same sense as the things*; rather, they hold (*gelten*) for the things. These classes, although something which holds, can now be envisaged as having a *second mode of being*. From them we can proceed, for example, to the cardinal numbers, which hold for these classes. … Cardinal numbers belong to a *third mode of being* and allow us to construct the fractions as relation extensions which hold for certain cardinal numbers . … These fractions can also be reified, that is, they can be envisaged as belonging to a *fourth mode of being*, and can be made elements of certain classes which hold for them, namely the real numbers. The latter belong to a *fifth mode of being*, while the complex numbers, being relation extensions that hold for certain real numbers, belong to a *sixth mode of being*, etc. (Carnap (1974), §42, emphases mine – KM)

Does Carnap take his talk of modes of being seriously? What does Frege's ex-student take their relation to existence or being as a second-order property to be? Answers are suggested by Carnap's account of the relation between his modes of being talk and his talk of spheres. In the first passage quoted above he seems to accept that objects belonging to different spheres enjoy different modes of being. But in a later passage from the *Aufbau*, he says that the constitution of an object on the basis of definite other objects not only does not mean that the object is of the same type (*gleichartig*) as the other objects but rather quite the opposite. If

> constitution (as is the case with spiritual (*geistigen*, mental) objects, in particular when they are of higher levels) leads to the formation of a new logical level, then *the constituted objects belong to a different mode of being, more exactly: to a new sphere of objects* (Carnap (1974), §151, emphases mine – KM; cf. §31)

No real friend of modes of being thinks objects "belong" (*gehören*), Carnap's term, to a mode of being If enjoying a mode of being is to be understood in terms of belonging to a sphere of objects, if the former way of talking is less "exact" than the latter, then Carnap's mode of being talk is mere loose talk, a way of talking which, as we have seen, he thinks drives home the great difference between objects belonging to one sphere and objects belonging to another sphere, a type of difference Husserl, as we have seen, refers to, in one case, as an "abyss", and which Carnap refers to as being "toto coelo different" (Carnap (1974), §31).

Whether or not Carnap's talk of modes of being really is loose talk, there is, we now see, the possibility that Husserl's references to and uses of modes of being should best be understood as loose talk about spheres or regions of being. When Husserl introduces his figurative or pictorial notion of a region he explains this notion very clearly in the following way: "highest essential generalities … delimit

regions or categories of individuals" (Husserl (1950), §2). In other words, the very idea of a region of being is understood in terms of the exemplification or instantiation of essences which are themselves taken to occupy nodes in trees, genus-species trees and determinable-determinate trees. As far as I can tell, Husserl employs the concepts of a *sphere* and a *domain* of being in a similar way. They are all understood in terms of the exemplification or instantiation of essence and not in terms of modes of being.[7]

It seems to me to be more likely that Carnap's talk of modes of beng is loose talk than that the same is true of Husserl's talk. But it is by no means impossible that Husserl who, throughout his career, endorses something like Bolzano's account of existence, did not really believe in modes *of* being. He talks, apparently indifferently, of correlations between acts and modes of being and of correlations between acts and regions. And it is difficult to believe that a philosopher who, like Husserl, takes great pains to *extend* Bolzano's account of existence to deal with propositions about what is real and what is merely fictitious and to *complement* it with an account of existence-in a domain, would never have considered the relation between Bolzano's account and modes of being, if he really took modes of being seriously.

Carnap's use of the language of modes of being is, as we have seen, designed to emphasize the phenomenon of difference between spheres. He thinks that "the difference of spheres of objects" is what "the frequently emphasized difference in recent philosophy between what is (*dem Seienden*) and what holds or is valid" (*dem Geltenden*) aims to bring out; the terminology of holding or validity, made popular by Lotze, is also, as we have seen, employed by Husserl when in 1894 he denies not only that there are modes of existence but also that there are modes of validity or holding. And Carnap's own distinction between objects and quasi-objects, he says, has the same aim. But in Carnap's multiplication of modes of being in the "stepwise progress of construction" quoted above, "in which the relationship between being and holding recurs several times", belonging to a new mode of being is to be understood, as we have seen, in terms of belonging to a new sphere of objects.[8]

The concepts of region, sphere, domain and boundary to be found in the writings of Husserl, Scheler, and Carnap are used to divide up or anatomise being. A different classification, orthogonal to those in terms of regions and spheres,

7 The difference between Husserl's regions and spheres seems to be that regions are, in the first place, regions of temporal items, physical, psychological, living, social individuals. A sphere is not so restricted; there is a sphere of ideality, of ideal objects.

8 All quotations in this paragraph are from Carnap (1974), §42.

and employed by James, Scheler, Husserl and vestigially by Wittgenstein, distinguishes between the world (microcosm, *Umwelt*) of a person or subject, on the one hand, and *the* world (the macrocosm), on the other hand.

In the list above of the modes of being endorsed by some of Brentano heirs, particularly by the phenomenologists, I attributed to Heidegger endorsement of four modes of being – the *Existenz* of *Dasein*, the being ready to hand of tools and being present to hand and life (Heidegger, 1972 (1927), p. 50). But in fact Heidegger's multiplication of modes of being in 1927 goes much further than this. Understanding, choosing and asking are modes of being (*Seinsmodi*) of questioners (Heidegger, 1972 (1927), §2). Authenticity and inauthenticity are modes of being of *Dasein* Heidegger, 1972 (1927), §9). Understanding of being is a determination of being (*Seinsbestimmtheit*) of *Dasein*. In the face of Heidegger's multiplication of modes of being or of close relatives thereof (determinations of being), one may wonder whether his 1927 philosophy of being is not simply the result of taking all too seriously Husserl's loose talk of modes of being and ignorance of second-order views. Heidegger's deathly serious ontology would then owe its origins to a failure to read with sufficient care Husserl's *Logical Investigations* or Frege.

The multiplication of modes of being by Heidegger and other phenomenologists raises an interesting question. When does multiplication become inflation? Where should the multiplier of modes stop? If asking is a mode of being, why not answering, baptising, declaring, ordering and promising? Does snow enjoy a mode of being which is peculiar to snow? If not, why not?

3 Correlates & correlationism

In the writings of many phenomenologists, there is the suggestion that there *is* a constraint on the multiplication of modes of being. They seem to have thought that something is a mode of being only if it stands in some sort of intentional relation, in a wide sense of the word, to (what they did not all call) a type of mental act or state. This is by no means the only or the earliest correlationist thesis in phenomenology. Husserl writes:

> The first breakthrough to this universal correlation a priori (*Korrelationsapriori*) of *object* of experience and *ways of givenness* (... around 1898) shook me so thoroughly that since then the entire work of my life has been dominated by the systematic development of the correlation a priori (Husserl (1976), 169 note).

Here the second term of the correlation is objects. Reinach makes a similar but more general claim: "To every objective something (*Gegenständlichen*) and class

of objects certain acts are correlated (Reinach (1989), p.483). And, as his examples make clear, he understands here by "object" both the bearers of properties but also certain properties. He does not however think that all mental acts and states have correlates: "it lies in the essence of moods (*Stimmungen*) that they do not require intentional correlates" Reinach (1989), p. 439), a neighbour of a view Heidegger was later to make famous. One early formulation of correlationism, in which two types of *knowledge* are correlated with two modes of being, is due to Meinong in 1910:

> There is perhaps no point of view from which the difference between existence and obtaining ... shows up more strikingly than in the modes of knowing (*Erkenntnisweisen*) naturally coordinated with these two kinds of being (*Seinsarten*). ... Existence is in principle known (*erkannt*) empirically, obtaining is in principle known a priori (Meinong (1983), p. 61 – tr. modified, Meinong (1977), p. 77).

Two types of knowledge which play a similar but not identical role are Husserl's (visual, tactile ... inner) perception and so called categorial perception. In 1913, Scheler formulates, perhaps for the first time, the idea that there is an "essential connection between forms of acts and forms of existence (*Daseinsform*)" (Scheler (1995), p. 232). In the same year, Husserl, although he allows for only a handful of modes of being, agrees: "Every mode of being... has according to its essence its modes of givenness ... ." (Husserl (1950), §77). Scheler employs the wretched expression "ontologism" or "absolute ontologism" for the denial of correlationism,

> the view that there could be objects which, in virtue of their essence, cannot be grasped by any consciousness. Every assertion of the existence of a kind of object requires, on the basis of this essential connexion, the specification of a type of experience in which this kind of object is given.[9]

This formulation, it should be noted, refers to a correlation between grasping or experience and kinds of objects, not modes of being. Absolute ontologism is the rejection of correlationism, a crime of which Scheler thinks Hartmann is guilty. Friends of correlationism, he implies, come in two main kinds. There are those who think of the correlates as interdependent – the view of Scheler and other "realist" phenomenologists – and those, like Kant, who think of the laws of objects as taking their "direction" from the laws for grasping objects and so favour a unilateral connection.[10] Przywara says that the "phenomenological principle of the *reciprocal*, inner conditioning of act and object" is what Scholasticism had

9 Scheler (1966), 270.

10 Scheler (1966), pp. 270 and 21.

formulated in the "axiom: *actus specificatur ab obiecto*"[11]. Whether any early realist phenomenologists thought that acts are specified by their correlates or that neither acts nor correlates enjoy any priority, Husserl, as Ingarden points out, thinks that "the difference in mode of givenness is what first brings with it that there is a difference in mode of being"[12]. It is *because* reality and consciousness are given in different ways that they enjoy different modes of being.

What, now, are the different correlations envisaged by the phenomenologists? What is the nature of the correlation? Amongst the answers to the first question are:

List 2:

– Perception	Being real/actual (Husserl)
– Resistance	Being real/actual (Scheler)
– Care	Being real (Heidegger)
– *Besorgen* (concern)	*Zuhandensein* (Heidegger)
– *Leibbewusstsein*	Being alive (Scheler)
– *Angst*	*In-der-Welt-sein* (Heidegger)
– Feeling	Being valuable (Scheler)
– Perception	Being present (Scheler)
– Memory	Being past (Scheler)
– Expectation	Being future (Scheler, Husserl)
– Conjecture	Being probable (Husserl)
– Phantasy	Being possible (Husserl, Scheler)
– Collective intentionality	Personal being (Scheler)

What is the nature of the correlation? Two distinct answers to this question can be distinguished in the writings of the philosophers referred to. According to the first answer, a mode of being is what acts of a certain type are intentionally related to or about. According to the second answer, a mode of being belongs to the correctness condition of a certain type of acts. Thus according to the first answer

(1) *a* is better than *b*

is the object of the act of preferring *a* over *b*, what this act is "intentionally directed towards". According to the second answer, (1) is the correctness condition for preferring *a* over *b* but the objects of this act are only *a*, *b*, in that order. According to both answers, being better than is to be understood as a mode of being, the

11 Przywara (1923), p. 72.

12 Ingarden (1964), p. 12. Cf. Scheler's description of Kantian correlationism above.

axiological mode. A variant of the first answer has it that, in certain conditions, preferring *a* over *b* constitutes knowledge of betterness. One popular candidate for this condition is correctness. Thus on this view, fear of a dog which is in fact dangerous amounts to a disclosure or revelation of the being dangerous or being disvaluable of the dog. Meinong, as we have seen, thinks that two distinct types of knowledge are correlated with the two modes of being he calls existence and obtaining. Similarly, Husserl writes:

> Every mode of being … has essentially its ways of givenness and so its own ways as regards methods of knowledge (Husserl (1950), p. §79)

Heidegger's correlationism often takes the form of the first answer: a mode of being is said to be what a certain type of acts or state is about, discloses or reveals. Two peculiarities of Heidegger's correlationism are that he has no time for either correctness conditions or for values, whether or not these are taken to be modes of being. In the case of the mode of being of being present-at-hand, he says that there is a type of "looking at something" (*Hinsehen*), in which there is "a holding back from any manipulation or utilization", and in which "the *perception* (*Vernehmen*) of the present-at-hand is accomplished" (Heidegger, 1972 (1927), §13). Being-in-the-world is "that in the face of which" and "for the sake of which *Angst* is anxious" (Heidegger, 1972 (1927), §40). The mode of being of reality "is referred back to the phenomenon of care" (Heidegger, 1972 (1927), §43(c)).

The varieties of correlationism considered so far correlate distinct types of modes of being with distinct mental acts or states. Another variety of correlationism distinguishes within one type of mental act or state distinct sub-types, each of which is correlated with distinct modes of being. This variety of correlationism is attributed to James by Marty. As we have seen, the Brentano-Marty account of being or existence essentially involves reference to the mental acts of accepting and rejecting. Thus one of Marty's arguments against modes of being is that "only if there is accepting and rejecting in different senses can being … and … nonbeing have different senses" (Marty, 1976 (1908), p. 324). And although Marty finds no reason to think that there is accepting and rejecting in different senses, he thinks that James makes just this mistake, in the opening salvo of his campaign against modes of being, in a long 1892 review of James's *Psychology*:

> If we look at what, according to our author, the content of the much mistreated concept [of existence] is supposed to be, one notices immediately that he continually confuses it with the very different concept of reality. Reality and existence signify according to him quite promiscuously sometimes what exists (i.e. correctly understood: everything that is correctly accepted, in contrast to the false), sometimes the real (i.e. the objective [*Sachhaltige*] in opposition to something which is a mere lack, a mere possibility, a mere phantasy etc.). The

> immediate consequence of this is that he takes the differences between what is real among themselves to be differences in the *mode of existence* and also takes the difference between something's being correctly acceptable and real, on the one hand, and correctly acceptable and non-real (e.g. a horse or a merely presented horse), on the other hand, to be differences in the *mode of existence*. This is a view he shares with many others. What is peculiar to James is this: he not only confuses the concepts of real and existing with one another, he also mistakes it for a third and a fourth thing. For he also calls real and existing everything which is as a matter of fact – even if in completely unjustified fashion – accepted by someone, as long as it is the object of a taking to be true. … In this way, he obtains new and in the end countless "ways of existence" (Marty, 1916 (1892), pp. 117–118).

Marty then notes that James multiplies not only ways of existence but *worlds*, a type of multiplication which was also to become very popular in twentieth century philosophy, in both CP and AP, from Scheler to Goodman:

> It turns out that according to him there are as many worlds, "each with its own special and separate style of existence", as one can distinguish classes of belief and illusion and of what is believed and of the illusory, from the descriptive and genetic points of view, both more general and more specific differences and indeed the most special differences realised only in one individual; from the "worlds" of the *idola tribus* and those of the religious belief common to entire ages and peoples to the uncountable worlds of individual opinions and madness. Everything is "real", but each merely "after his own fashion". Interest, which James identifies with belief, also contributes with all its differences (aesthetic, practical etc …) to the multiplication of "modes of existence". And since the author rightly distinguishes degrees of intensity in the case of interest and, with rather less justification, in the case of belief, he then begins to speak in all seriousness of a difference between what is more or less "real" or existing. A beginner faced with this bundle of equivocations and the subjectivism bound up with it, which muddles all concepts, would be completely perplexed (Marty, 1916 (1892), p. 118).

The phenomenologist to whom, as already noted, we owe the most careful account of modes of being, Ingarden, in his 1948/9 masterpiece, *Spór o istnienie świata*, is also an exception to many of the generalisations in this section. He is dismissive about earlier attempts to understand modes of being (presumably those of Scheler and Heidegger) and says of his own work that it seems to him to be the "first attempt to go beyond the vague locutions in which the different ways of being are normally spoken about and to replace them by rigorous concepts (Ingarden (1964), p. 257). Ingarden's modes of being differ from those of other phenomenologists in two respects. First, although, as already noted, a friend of correlationism, Ingarden indicates no mental acts or states as the correlates of his three most basic modes of being, the enduring of substances (such as a chair), the going on of processes (such as a deliberation) and the occurring of punctual events (such as judging, deciding or winning a race). Perhaps the most plausible candidates are the perception of things and the perception of processes, two types

of perception often distinguished by psychologists. Second, his modes of being are analysed into different factors, makers or "existential moments". These are arguably all formal – four different types of dependence and independence (*Seinsautonomie* vs *Seinsheteronomie*, *Seinsursprunglichkeit* vs *Seinsabgeleitetheit*, *Seinsselbständigkeit* vs *Seinsunselbständigkeit*, *Seinsabhängigkeit* vs *Seinsunabhängigkeit*), in a wide sense of the word, and yield over a dozen modes of being. If Ingarden is right, modes of being, like Bolzano-Frege existence, are formal or topic-neutral.[13]

Ambitious correlationism, then, is a distinguishing feature of phenomenology. Within AP *some* such correlations have been popular. Between phantasy, imagination or conceiving, on the one hand, and possibility, on the other hand and between emotions and value, as in buck-passing or fitting attitude accounts of value. Kenny's influential account of the formal objects of mental acts is, as he notes, a development of the already mentioned scholastic axiom "actus specificatur ab obiecto"[14] which, as we have also already noted, is the precursor of the correlationisms of Brentano and his heirs. Perhaps the most influential forms of correlationism within AP have been verificationisms about meaning and about the truth (obtaining) of truth-bearers (states of affairs), verificationisms which go back to Husserl.[15]

4 Essence vs modes of being

If an object enjoys some mode of being, it does so in virtue of its nature or essence. As we have noted in §2, this a thesis accepted by Meinong. It is also accepted by the early phenomenologists:

> A mode of being is determined by (*richtet sich nach*) essence (Scheler (1995), p. 296, cf. p. 285).

13 On Ingarden's modes of being, cf. Haefliger (1994), Simons (2012), Richard (2017).

14 Kenny (1963), p. 189.

15 Cf. Mulligan (2017). The view, referred to above, that different types of knowledge are correlated with different modes of being is entailed by Husserl's verificationism and commitment to modes of being. Although Scheler never gave up on the correlation between forms of acts and forms of being he did reject Husserl's verificationism, like Weyl. For reasons of space many of the claims in this section are not documented. For some details about correlationism, cf.Mulligan (2017a).

> Which way of being is it which is peculiar to an object, which is therefore determined (*vorbestimmt*) by its essence …? (Ingarden (1964), p. 58)

They also thought, following Husserl, that essence grounds necessity and necessary possibilities, whether or not they conceived of being necessary as a mode of being. Similarly, as we have seen, they often say that the correlations between acts and modes of being are themselves rooted in essences.

The unanimity about the priority of essence came to an end when Heidegger (following Scheler's discussion of the question whether man, as opposed to person, has an essence) declared in 1927:

> The "essence" of this being [*Dasein*] lies in its to-be (*Zu-sein*). Its what-being (*essentia*), must, to the extent that one can speak of it at all, be understood in terms of its being (*existentia*) (Heidegger, 1972 (1927), §9).

But Heidegger, unlike Sartre a few years later, sometimes forgets himself. Thus he says:

> If to *Dasein* there belongs essentially the mode of being of Being-in-the-world, then … (Heidegger, 1972 (1927), §18).

> And because *Dasein* is essentially in each case its possibility, … (Heidegger, 1972 (1927), §9).

How might one distinguish between essences and modes of being? Modes of being, their friends believe, are not properties, neither first-order nor second-order properties. Perhaps modes of beings are what different restricted quantifiers represent. This is the route followed by McDaniel: different modes of being are canonically represented by different semantic, primitively restricted quantifiers. This is not the route followed by the phenomenologists. I suspect that they were led to modes of being by the following considerations. Like many earlier philosophers, they admitted non-substantial particulars. One type of such particulars is what have been called individual accidents, tropes, particularised properties. A second type of non-substantial particular is what has often been called individual essence. A friend of the first type of non-substantial particulars allows for the possibility that Sam's sadness at t is numerically different from but completely similar to Maria's sadness at t. A friend of the second type of non-substantial particulars says that Sam's humanity is numerically different from Maria's humanity and that they are completely similar. Sam's humanity cannot fail to resemble Maria's humanity. But the sadness of one person at a time may resemble less than completely the sadness of another person at a time. Sam's humanity is as particular and nonrepeatable as his sadness. Sam cannot exist unless his humanity exists but can exist at a time even though nothing at the time makes it true that he is sad. The

phenomenologists sometimes mark the difference between individual essences and general essences by reserving the term "*Wesen*" (essence) for the former and "*Wesenheit*" (essentiality) or "*Idee*" for the latter.

Suppose, then, that you allow for the *particularisation* of general properties and of essences and ask whether being real, being physical, being alive, enduring etc. can be particularised. If Sam is alive, is there an individual accident which is his being alive, his living? There is his life, which is numerically distinct from Maria's life. But these two lives are longish processes. Clearly, predications of the form "*x* is alive" are not made true by distinct vitality tropes. Is being alive an individual essence? No. It is a form of adjectival being (as living is a form of verbal being). Nor is it a general essence. If being alive is a general property, it is a general property which allows of no particularisations, of no corresponding tropes or individual essences. A philosopher who asks all these questions and answers them in the ways mentioned might well conclude that being alive differs so much from general properties that it deserves an ontological category of its own – modes of being. A more plausible view is that what have been called modes of being are in fact good old properties and relations but properties and relations which are distinguished by the fact that they *admit of no particularisations*.

For those of us who think that what have been called modes of being are in fact properties and relations, the claim that such properties and relations correspond to no non-substantial particulars, neither to tropes nor individual essences, provides a second way of circumscribing such properties and relations in addition to correlationism.

Why should modes of being be rejected? If existence is a second-order or a first-order property, it is a formal property. Formal properties do not have material determinates[16] and most of the candidates for the rôle of modes of being mentioned above are not formal. One objection to the view that existence is a second order property is that, on both the Bolzanian and Fregean versions of this view, Platonism is presupposed. One may well think that an account of existence should not presuppose Platonism or indeed nominalism. If existence is understood, following Brentano and Marty, in terms of what it is correct to accept or the knowability of this,[17] Platonism is avoided. But there is a fatal objection to Brentano's view, which is suggested by Husserl's rhetorical question about it: "Is the concept of a judgment really supposed to be included in the concept of existence as a compon-

16 According to Wittgenstein, formal concepts have no determinates, According to Husserl, they have no material determinates. Cf. Mulligan (2013).

17 Kriegel (2018), ch. 5, defends Brentano's view of existence against many rivals; Textor (2017) defends a neo-Brentanian account of existence.

ent?".[18] Textor notes that Moore makes a related objection to Brentano on truth, deploying one of his his open-question arguments:

> that it is false appears to be plain from the fact that we can raise the question whether it *is* right to believe everything that is true: that is to say, we are immediately aware that true and rightly believed are two distinct concepts, one of which, true, is an unanalyzable property belonging to some objects of belief.[19]

Husserl rightly thinks that the concepts of truth and existence are simple, logical[20] and so formal. If the concept of existence contained that of judgment or belief as a component, it would be a non-formal concept. It is perhaps because he thinks of the concept of existence as a material concept that Marty is prepared to allow that existence may have modal determinates. Is the commitment to Platonism of the Bolzano-Frege-Russell-Husserl view, on the one hand, more costly than the mistake of taking existence to be a non-formal notion? For present purposes, it is enough to note that if either of our two Austrian priests is right, modes of being of objects are superfluous to requirements.[21]

This paper was prepared as part of the Swiss FNS projects, *Connectives, Predicates and Priority* and *The Nature of Existence*. It is based on courses in Geneva and Lugano and on material presented to audiences in Matera (2015), Warsaw (2017), Oxford and Lugano (2018). I am grateful to these audiences for their comments. I am also very grateful to Hamid Taieb for showing me his incisive, forthcoming paper on modes of being in early phenomenology and to Erwin Tegtmeier, Kris McDaniel and Denis Seron for stimulating exchanges and help.

Bibliography

Ameseder, R. (1904), "Beiträge zur Grundlegung der Gegenstandstheorie", in *Untersuchungen zur Gegenstandstheorie und Psychologie*, edited by A. Meinong, Leipzig: Barth, 51–120.

18 Husserl (2001), p. 218.

19 Moore (1903), Textor (2017).

20 Husserl (2001), p. 215.

21 The rejection of modes of being for objects does not by itself amount to a rejection of modes of being for states of affairs (cf. footnote 2 above). Suppose there are states of affairs and that (following Prior) connectives are not predicates or relational expressions or notational variants thereof. What, then, are the semantic values of connectives, if they have such values? One not implausible answer is that they are not properties or relations. At this point, a friend of modes of being may see an opening: probability, value, past, present and future and temporal priority, she might suggest, are the modes of being of states of affairs.

Bergmann, G. (2004(1967)), *Realism. A Critique of Brentano and Meinong*, Frankfurt: Ontos Verlag.
Bolzano, B. (1929), *Wissenschaftslehre*, Leipzig: Meiner.
Brentano, F. (1978), *Grundlegung und Aufbau der Ethik*, Hamburg: Meiner.
Carnap, R. (1974), *Der logische Aufbau der Welt*, Frankfurt am Main: Ullstein.
Grossmann, R. (1984), *Phenomenology & Existentialism: An Introduction*, London: Routledge.
Haefliger, G. (1994), *Über Existenz: die Ontologie Roman Ingardens*, Dordrecht : Kluwer.
Heidegger, M. (1972(1927)), *Sein und Zeit*, Tübingen: Niemeyer.
Husserl, E. (1950), "Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie I ", in *Husserliana* III, edited by W. Biemel, The Hague: Nijhoff.
Husserl, E. (1964), "Erste Philosophie", in *Husserliana* VIII, The Hague: Nijhoff.
Husserl, E. (1975), "Logische Untersuchungen", in *Husserliana* XVIII, *Prolegomena zur reinen Logik*, The Hague: Nijhoff. English transl.: "Logical Investigations", translated by J. Findlay, London 1973: Routledge and Kegan Paul.
Husserl, E. (1976), "Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie. Eine Einleitung in die phänomenologische Philosophie", in *Husserliana* VI, edited by J. W. Biemel, The Hague: Nijhoff.
Husserl, E. (1979), "Intentionale Gegenstände", in *Aufsätze und Rezensionen*, edited by B. Rang, The Hague: Nijhoff, 303–348.
Husserl, E. (1984), *Logische Untersuchungen*, Vol. II, first part, The Hague: M. Nijhoff.
Husserl, E. (2001), *Logik*: *Vorlesung* 1896, Dordrecht: Springer.
Husserl, E. (2009), *Untersuchungen zur Urteilstheorie. Texte aus dem Nachlass* (1893-1918), edited by R. D. Rollinger, The Hague: Nijhoff.
Ingarden, R. (1948/49), *Spór o istnienie świata* [*The Controversy over the Existence of the World*], Cracow: Polska Akademia Umiejetności, I 1947, II 1948, III 1961. German translation *Der Streit um die Existenz der Welt*, Tübingen: Niemeyer, I, II/1, II/2 1964, III 1961. Partial English translation *Time and Modes of Being*, translated by H. R. Michejda, Springfield: Charles Thomas, 1964.
Ingarden, R. (1964), *Der Streit um die Existenz der Welt*, Bd. 1, in *Existentialontologie*, Tübingen: Niemeyer.
Ingarden, R. (1965), *Der Streit um die Existenz der Welt*, Bd. 2, in *Formalontologie*, Teil Form und Wesen, Tübingen: Niemeyer.
Ingarden, R. (1981), *Ontology, Identity, and Modality*, Cambridge: Cambridge University Press.
Ingarden, R. (2014), *Existence*: *Essays in Ontology*, Cambridge: Cambridge University Press
James, W. (1952(1890)), *The Principles of Psychology*, Chicago: The University of Chicago.
Kenny, A. (1963), *Action, Emotion and Will*, London: Routledge.
Kriegel, U. (2018), *Brentano's Philosophical System. Mind, Being, Value*, Oxford: Oxford University Press.
Marty, A. (1916(1892)), "Anzeige von William James' Werk: The Principles of Psychology", in *Gesammelte Schriften*, I 1, edited by J. Eisenmeier, A. Kastil & O. Kraus, Halle: Niemeyer, 105–156.
Marty, A. (1976(1908)), *Untersuchungen zur Grundlegung einer allgemeinen Grammatik und Sprachphilosophie*, Vol. I (only volume published), Hildesheim/New York: G. Olms.
Meinong, A. (1977), *Über Annahmen*, Gesamtausgabe, Bd. IV, edited by R. Haller, Graz: Akademische Druck- u. Verlagsanstalt.
Meinong, A. 1983), *On Assumptions*, translation of Meinong (1977), translated by J. Heanue, Berkeley: University of California Press.

McDaniel, K. (2017), *The Fragmentation of Being*, Oxford: Oxford University Press.
Moore, G. E. (1903), "Review of Brentano 1902 *The Origin of the Knowledge of Right and Wrong*, a translation of *The Origin of the Knowledge of Right and Wrong*", in *International Journal of Ethics*: 14, 115–123. Reprinted in *The Philosophy of Brentano*, edited by L. McAlister, London: Duckworth, 176–181.
Mulligan, K. (2013), "Formal Concepts", in *Studies in the History and Philosophy of Polish Logic. Essays in Honour of Jan Woleński*, edited by K. Mulligan, K. Kijania-Placek & T. Placek, Palgrave Macmillan, 205–223.
Mulligan, K. (2017), "Brentano's Knowledge, Austrian Verificationisms, and Epistemic Accounts of Truth and Value", in *Monist*: 100(1), special number on Brentano, Guest editor, U. Kriegel, 88–105.
Mulligan, K. (2017a), "Incorrect Emotions in Ancient, Austrian & Contemporary Philosophy", in *Revue de philosophie de la France et de l'étranger*: 100(1), special number on Brentano, Guest editor, G. Fréchette, 491–512.
Mulligan, K. (2019), ""Misleading Pictures, Temptations & Meta-Philosophies: from Marty to Wittgenstein", in *Marty and Contemporary Philosophy*, edited by G. Bacigalupo and H. Leblanc, Palgrave Macmillan, 197–232.
Przywara, E. (1923), *Religionsbegründung. Max Scheler – J. H. Newman*, Herder: Freiburg.
Reinach, A. (1989), *Werke*, Vol. 1 of Sämtliche Werke. Textkritische Ausgabe in 2 Bänden, edited by Karl Schuhmann & Barry Smith, Munich : Philosophia Verlag.
Richard, S. (2017), "Catégories d'objet et modes d'être chez Meinong", in *Les Études philosophiques*: 173(3), 367–384.
Scheler, M. (1955), "Vom Umsturz der Werte", in *Gesammelte Werke*, III, Berne: Francke.
Scheler, M. (1966), *Der Formalismus in der Ethik und die materiale Wertethik. Neuer Versuch der Grundlegung eines ethischen Personalismus*, in *Gesammelte Werke*, II, Berne: Francke.
Scheler, M. (1995), *Späte Schriften*, in *Gesammelte Werke*, IX, Berne: Francke.
Simons, P. (2012), "Four Categories – and More", in *Contemporary Aristotelian Metaphysics*, edited by T. Tahko, Cambridge: Cambridge University Press, 126–139.
Specht, E. K. (1967), *Sprache und Sein*, Berlin: de Gruyter.
Taieb, H. (forthcoming), "A Paleo-Criticism of Modes of Being: Brentano and Marty".
Textor, M. (2017), "Towards a Neo-Brentanian Theory of Existence", in *Philosophers' Imprint*: 17(6).
Vallicella, W. F. (2002), *A Paradigm Theory of Existence. Onto-Theology Vindicated*, Dordrecht : Kluwer.
Van Inwagen, P. (1981), *Ontology, Identity, and Modality*, Cambridge: Cambridge University Press.
Van Inwagen, P. (2014), *Existence: Essays in Ontology*, Cambridge: Cambridge University Press.

Takashi Yagisawa

Imagining Fictional Characters

Abstract: I defend an argument against the view that fictional objects are abstract objects manufactured by human beings. The key move in the argument is that since abstract objects are not subject to perception, they are not subject to imagination. I bolster the argument by defending the importance of imagination in our appreciation of fiction and by critically examining some attempts to skirt the key move. In doing so, I highlight the difference between imagination *de re* and imagination *de dicto* as crucial.

1 Introduction

Abstract Artifact Theory (AAT) is the metaphysical theory which says that all fictional objects are abstract objects, actually created artificially by authors and kept in existence by consumers of fiction.

I wish to object to AAT by pointing out a certain difficulty caused by the abstractness AAT attributes to fictional objects. The difficulty concerns the fact of our imaginative engagement with fictional objects in general, and fictional characters in particular. When we read a work of ordinary narrative fiction, we imagine the fictional characters that appear in the fiction. Our imagining the fictional characters is almost always vital to our full appreciation of the work as literature.

Such imaginative engagement with fictional characters may precede and even be required by genuine emotional response to the contents of the fictional story,[1] but it is important to keep them separate. Having an emotional response to the story's contents may raise some issues for the defenders of some theories about the nature of narrative fiction, but my argument against AAT is independent of such issues. For my argument against AAT, taking note of the metaphysical nature of fictional characters is enough, and considerations going beyond it are irrelevant.

0 This work was supported by California State University Northridge, College of Humanities Faculty Fellowship and Grant Program for the Spring 2017.

1 See Walton (1978), pp. 195–204 and Ch. 7.

https://doi.org/10.1515/9783110664812-011

2 The argument

Here is my argument against AAT:

1. We sometimes imagine some fictional characters.
2. We cannot imagine abstract artifacts of the kind postulated by AAT.

Therefore,

3. Some fictional characters are not abstract artifacts of the kind postulated by AAT.[2]

The argument is valid, but the two premises are open to debate. The defender of AAT would undoubtedly deny one or both of them. Let us begin our consideration with premise 1.

2.1 Premise 1

Premise 1 is an understatement as an observation concerning proper literary experience. When reading narrative fiction, it is important that the reader exercise the ability of imagination to follow the story. Imaginative engagement with the fictional characters in the story is a vital part of the literary experience of reading a narrative fictional story and contributes to enhancement of the literary appreciation of the story.[3] It is possible to read a narrative fictional story without imagining its fictional characters, but doing so would almost certainly leave the reader without important aspects of literary experience. Not imagining Frankenstein's monster as grotesque, the hunchback of Notre Dame as visibly disfigured, and Anne of Green Gables as a spirited freckled redhead would rob the reader of the full experience of the respective stories in which these fictional characters appear.[4]

In order to experience a narrative fictional story fully as it should be experienced as literature, we should imagine all (key) fictional characters in the story. But we do not always do what we should, and we in fact imagine only some (key)

2 I sympathize with a similar line of objection against AAT in Sainsbury (2009), pp. 91–114.

3 Let me emphasize that it is imagining fictional characters that my discussion focuses on, not imagining ourselves (along with fictional characters) in the fictional setup. I have nothing to say about imagining ourselves in any setup. Cf. Walton (1997).

4 More good examples may be found in Isaac Asimov's novelette *Nightfall*. No particular fictional character (in the sense of a fictional person) needs to be the object of imagination, but the fictional planet Lagash and fictional goings-on near its surface need to be imagined vividly for the full appreciation of the novelette. For my purposes in this paper, the distinction between fictional characters and fictional objects (and goings-on) is unimportant.

fictional characters only some of the time. The above argument rests on what we do, instead of what we should do.

Premise 1 does not say that all of us imagine all fictional characters whenever we read the fictional stories which feature them. Some of us may lack sufficiently strong power of imagination of the adequate sort, and therefore may never imagine any fictional character. Those of us who possess sufficiently strong power of imagination of the adequate sort may sometimes fail, for whatever reason, to imagine some fictional characters appearing in fictional stories when we read the stories. But some of us do sometimes imagine some fictional characters when we read the fictional stories in which they appear. This is how premise 1 should be understood. This is good enough for my purposes against AAT, for AAT claims that all fictional characters are abstract (throughout their existence).

To focus our attention in deliberating the plausibility of premise 1, let us take a passage from an example work of fiction, Shakespeare's *Romeo and Juliet*, Act 1, Scene 5:

JULIET
Good pilgrim, you do wrong your hand too much,
Which mannerly devotion shows in this;
For saints have hands that pilgrims' hands do touch,
And palm to palm is holy palmers' kiss.

ROMEO
Have not saints lips, and holy palmers too?

JULIET
Ay, pilgrim, lips that they must use in prayer.

ROMEO
O, then, dear saint, let lips do what hands do;
They pray, grant thou, lest faith turn to despair.

JULIET
Saints do not move, though grant for prayers' sake.

ROMEO
Then move not, while my prayer's effect I take.

Kisses her

As we read this, we imagine, that is, picture a scene in which two fictional characters, Romeo and Juliet, interact with each other verbally and physically. The pictured scene includes in particular an occurrence of touching of their hands and

an occurrence of touching of their lips. Visually imagining such occurrences is important to the full appreciation of the play. Such visual imagining is imagining not just any occurrences but occurrences of touching of hands and touching of lips, and imagining not just any pair of entities but Romeo and Juliet. This is an example of imagining particular fictional characters as they appear in a fictional story engaging in spatiotemporal acts. Such imagining takes place typically as we read the relevant part of the fictional story, but it may also happen at other times, for example, when we think about what we read in the past.

The defender of AAT might deny premise 1 and claim that we never imagine fictional characters and that when we appear to imagine fictional characters, we are not really imagining anything at all, or else are really imagining some entities corresponding to the fictional characters, but never the fictional characters themselves.

It is true that those of us with sufficient imaginative power of the relevant kind may sometimes fail to exercise the power when reading a fictional story, either because we are distracted by other considerations concerning the story, or because we do not care enough about appreciating the story, or because of sheer imaginative laziness. But usually we do manage to imagine something as we read through the story, and typically our act of imagination involves imagining the fictional characters depicted in the story as they are depicted. In fact, in most narrative fiction it is hardly an exaggeration to say that one of the main functions of fictional characters is to be the objects of our imagination. To claim that we do imagine some things while reading a narrative fiction depicting fictional characters but what we imagine are some things other than the fictional characters would be to rob an important point of our speaking of fictional characters as existent entities. So the burden is on the defender of AAT to show what we are really imagining, or what we are really doing, when we seem to imagine the fictional characters.

The pretense theory of fiction[5] might prompt the defender of AAT to say that although we seem to imagine Romeo and Juliet, we really just pretend to imagine them. The problem of this line of thought is that the intelligibility of pretending to do X presupposes the intelligibility of doing *X* but that the intelligibility of imagining Romeo and Juliet as postulated by AAT is doubtful. Like imagining the number 17 or the Platonic Form of Beauty, imagining the artifactual abstract objects that are, according to AAT, Romeo and Juliet is hardly fathomable. (It may be noted here that in critiquing this move by the defender of AAT concerning Premise 1, I am touching on Premise 2, in effect defending a claim stronger than Premise 2 to the effect that imagining an abstract object is unfathomable.) Compare this with

5 See Evans (1982), Ch. 10; Kripke (2013); and Walton (1990).

clearly intelligible cases of pretending, e.g., pretending to serve a pie by pushing a wad of mud.

2.2 Premise 2

To see how plausible premise 2 is, we need to remember that imagining in the sense relevant to our discussion is primarily perceptual.

What distinguishes imagining from conceiving is that the former is essentially perceptual but the latter is not. Imagine a chimpanzee. Then imagine that the chimpanzee has a pipe. Now, where is the pipe? The answer is "In the chimpanzee's mouth", or perhaps "In the chimpanzee's hand". As we answer the question this way, we focus on the visual image of the pipe in the mouth or hand. If we are instead to conceive a chimpanzee having a pipe, we may well conceive a chimpanzee having a pipe in its mouth or hand, but we do not have to. We may just as well conceive a chimpanzee having a pipe without conceiving the chimpanzee having the pipe in any particular location or in any particular way. Conceiving is essentially only conceptual but imagining is essentially perceptual as well as conceptual.

In the case of the example of Romeo and Juliet, the imagination in question is auditory (hearing the words spoken by Juliet and Romeo), visual (seeing their bodily movements), and perhaps tactile (vicariously feeling the touching hand/lips).[6] If we are disallowed to use any perceptual mode of imagination, we will be unable to imagine Romeo or Juliet. Not only does perception take place in spacetime but perception is spatiotemporal by nature in the sense that it has to be of some thing or occurrence that is spatiotemporal if it is to be of anything at all. Because of this, imagining is spatiotemporal by nature in the same sense. That is, imagination (in the relevant sense) inherits spatiotemporal nature from perception.

Cast as an argument specifically about Romeo and Juliet in this Scene, the argument against AAT goes as follows:

4. We do imagine Romeo kissing Juliet, and when we do, we are imagining the fictional character Romeo kissing the fictional character Juliet.
5. It is impossible to imagine an abstract object kissing an abstract object.

6 Possibly also olfactory, and even gustatory.

Therefore,

6. The fictional characters Romeo and Juliet are not abstract objects.

Again, this argument is valid but its premises are open to debate. Take the second premise. Imagining in the pertinent sense is perceptual, and in particular visual. So, to imagine *x* kissing *y* is to visually imagine, that is, picture *x* as kissing *y*. But since abstract objects completely lack visually accessible characteristics, we cannot visually imagine (picture) abstract objects as doing anything spatiotemporal. So, we cannot imagine an abstract object kissing an abstract object.

Do not confuse this reasoning in support of the second premise with the following different reasoning in support of the second premise: (i) abstract objects are not spatiotemporal; so (ii) they cannot kiss anything (because kissing is a spatiotemporal activity); so, (iii) they cannot be imagined as kissing anything, including kissing an abstract object. (A similar argument may be given by speaking of being kissed instead of kissing.) The reasoning "(i), (ii), so (iii)" is vulnerable in a way the above reasoning is not. Someone might maintain that the impossibility of *x* kissing *y* does not entail the impossibility of our imagining *x* kissing *y*, for she might maintain that it is doubtful that we cannot imagine anything impossible. Such a person might say that it is possible for us to imagine Hesperus not being identical with Phosphorus, or water not being H_2O, or lightening not being a stream of electrons, all of which are impossible. It is true that Saul Kripke has a powerful response against the claim that we can conceive, let alone imagine, failures of these identities, by means of the celebrated idea of an "epistemic counterpart".[7]

But Kripke's position might not be inevitable, and the contrary position might be maintained consistently. My argument against AAT does not rest on the presumption that the position contrary to Kripke's could not be maintained consistently. In particular, my argument does not rely on the assumption that impossibilities are unimaginable. It only uses the assumption that non-spatiotemporal objects are unimaginable, or more narrowly, that non-spatiotemporal objects cannot be imagined to be kissing anything (or that non-spatiotemporal objects cannot be imagined to be being kissed by anything). This assumption is more plausible than the assumption that impossibilities are unimaginable.

7 Kripke (1980), pp. 97-105 and 116-143.

3 Two kinds of predication

The defender of AAT might resort to a well-known distinction concerning predication in order to object to my argument. According to AAT, Romeo and Juliet are abstract objects, hence they do not have any properties that require concreteness, like kissing someone and being kissed by someone. In this sense, it is incorrect to predicate kissing of Romeo or being kissed of Juliet. At the same time, in the play Romeo is depicted as kissing Juliet, and Juliet as being kissed by Romeo. So in this sense it is correct to predicate kissing of Romeo and being kissed of Juliet. The two leading contemporary theorists who postulate such a distinction of predication both leave the distinguished kinds of predication primitive.[8] Fortunately, one kind of predication is familiar enough and needs no definition or any kind of characterization beyond presentation of a few typical examples. This is the kind of predication such that Shakespeare was human, Eurasia is a continent, and Betelgeuse is a star. Also the predication of the property of being fictional of Juliet is of this kind. But when we say, "Juliet was kissed", and say something true, our predication of the property of being kissed is of the second kind. In this second sense, it is not truth that Shakespeare was human, that Eurasia is a continent, or that Betelgeuse is a star unless these entities are fictional entities depicted as being these respective ways in some fiction. Let use adopt van Inwagen's terminology and use "have" and "hold" corresponding to the first and second kinds of predication, respectively.

The defender of AAT might try to use this have/hold distinction to defend AAT by saying that we can imaging Romeo kissing Juliet not in the sense of imagining Romeo and Juliet having the relation of kissing but in the sense of imagining Romeo and Juliet holding the relation of kissing. This defense of AAT suffers from two defects.

First, it does not do justice to the phenomenology of our imaging Romeo kissing Juliet. When we are imaginatively engaged in the story of *Romeo and Juliet*, and imagining Romeo kissing Juliet, we are not imagining Romeo kissing Juliet according to the story. I am not denying that we could imagine Romeo kissing Juliet according to the story, but doing so would not be the same as imagining Romeo kissing Juliet, and it is the latter that enables us to imaginatively engage ourselves with the story. The former simply situates us outside the story as it were while we

8 See Van Inwagen (1977) (predicating vs. ascribing), Van Inwagen (1983) (having vs. holding), and Zalta (1983) (exemplifying vs. encoding).

do our imagining, and detaches us from imaginative engagement with the story in the sense relevant to our discussion.

Second, suppose that we interpret this defense of AAT charitably and understand it as saying that concerning the properties and relations Romeo and Juliet hold, we imagine Romeo and Juliet as having these properties and standing in these relations. That is, suppose in particular that we understand the defense of AAT as claiming that (i) concerning the property of kissing (which Romeo holds), we imagine Romeo as having it, (ii) concerning the property of being kissed (which Juliet holds), we imagine Juliet as having it, and (iii) concerning the relation of kissing (which the pair Romeo-Juliet holds), we imagine the pair Romeo-Juliet as having.

Even under such a charitable supposition, this defense of AAT misses the mark. My objection does not concern how we imagine Romeo and Juliet as being. It concerns what individuals we imagine when we imagine Romeo and Juliet. That is, my objection is not the contents of our imagination so much as the objects of our imagination. My contention against AAT is that Romeo and Juliet, as conceived by AAT, cannot be the *res* of our imagination *de re*. Since we cannot imagine them, we cannot imagine them as being anyway at all.

4 Social constructs

The defender of AAT might draw an analogy with socially constructed institutions. Take a sports team as an example, say, the Los Angeles Dodgers. It might be said that we can perceive the Dodgers play against the San Francisco Giants despite the fact that the team Dodgers is an abstract artifact (as is the team Giants). We perceive the team by perceiving its member players on the field. We perceive the players literally and are thereby said to perceive the team, as it were, by courtesy. The players are the objects of literal perception, whereas the team is the object of perception by courtesy. Likewise, the defender of AAT might claim that the fictional characters, Romeo and Juliet, are objects of perception by courtesy, hence they can be objects of imagination by courtesy. If Romeo and Juliet are indeed objects of perception by courtesy, then the question is: What are the objects of literal perception by means of which Romeo and Juliet are said to be objects of perception by courtesy?

At this point the defender of AAT might choose to shift our attention to theater. When we attend theater to see the play *Romeo and Juliet*, we perceive, among others, two actors as they play Romeo and Juliet on stage. Call them "Bobby" and "Betty". Bobby and Betty are not fictional characters. They are *playing* the fictional

characters, Romeo and Juliet. It is clear that we perceive Bobby and Betty. And via the relation of playing, our perceptual – hence imaginative – *rapport* might be said to somehow extend from the playing actors to the played characters.

This picture might be expanded to apply to the case of reading the written text of the play. As we read the text, we may imagine attending theater and perceiving Bobby and Betty on stage. The defender of AAT would then remove the stage and the entire theatrical setup, situate Bobby and Betty against the background depicted in the play, and invite us to imagine Bobby kissing Betty against that background. (If we wished to begin with more realistic background than stage set, we might speak of watching a motion-picture version of the play in a movie theater.) This imagining of Bobby and Betty would be as straightforwardly perceptual as perceiving Bobby and Betty. If we should accept the invitation and imagine them, we would succeed in imagining the individuals playing the fictional characters, and in doing so we would succeed in imagining the fictional characters themselves by courtesy. Or so the defender of AAT might say. Would this be satisfactory?

I am afraid not. The problem is that imagining by courtesy is not imagining. When we are genuinely engaged in the play, we imagine Romeo and Juliet, not Bobby and Betty, even though we may be perceiving Bobby and Betty. Of course, we need not imagine Romeo and Juliet by perceiving or imagining Bobby and Betty. Noting this fact, someone might suggest that perhaps we may imagine Romeo and Juliet not by imagining particular actors but by imagining arbitrary actors playing Romeo and Juliet. But if we followed this suggestion, we would lose the sense of imagining *de re*. Imagining arbitrary actors playing Romeo and Juliet is indistinguishable from imaging that an actor looking thus-and-so behaving toward an actress looking thus-and-so. Note that the notion of playing-Romeo that is relevant here is understandable without reference to any particular res, either an actor or a fictional character; it is a purely *de dicto* notion to the effect of moving thus-and-so and saying such-and-such words.

Here is a further thought on the relation of playing that may count against AAT. The relation of playing is holistic; Bobby cannot play Hamlet without Betty (or someone else) playing Juliet, and Bobby and Betty cannot play Romeo and Juliet without others playing the supporting characters and the surrounding setup.[9] In view of this, the relation of playing had better be replaced with the holistic two-

9 Except that Bobby may "air play" Romeo and Bobby and Betty may "air play" Romeo and Juliet in the sense analogous to the sense of "air guitar," which is a guitar-playing movement without a guitar. But "air play"-ing a fictional character is not playing the fictional character any more than "air guitar"-ing is playing the guitar.

place relation of *bringing* or *presenting* between a theater troop and a world of fiction. This relation is *de dicto*; what the troop needs to do is purely *de dicto*, putting up a performance satisfying a certain purely qualitative description. If someone objects to this, let her compare one scenario in which some *de re* connections with abstract artifacts play a crucial role and another scenario in which no such connection exists; there would be no difference in what the troop members intentionally do.[10] So there is no *de re* basis to enable us to imaging *de re* about Romeo and Juliet by perceiving and/or imagining Bobby and Betty.

Back to the baseball analogy: Do Bobby and Betty stand in an analogous relation to Romeo and Juliet as the Dodgers players stand to the team Dodgers? No. Bobby and Betty do not make up Romeo and Juliet the way the players (along with the general manager, the coaches, and the owner) make up the baseball team. The relation between the actors and the fictional characters is not even relevantly analogous to the relation between the players and the team. In an important sense, there would be no team without players, but there would still be the fictional characters without actors. Nobody needs to "play" Romeo or Juliet in order for Romeo and Juliet to exist as abstract fictional characters. We may be said to be able to perceive by courtesy, or perhaps even literally, what Romeo and Juliet did according to Shakespeare's play by perceiving Bobby and Juliet enact the relevant scene. But that still does not imply that we thereby perceive the abstract fictional characters in the relevant *de re* sense. No perception, no imagination.

The defender of AAT might say that when we imagine Bobby and Betty play Romeo and Juliet, we are really directly (rather than by courtesy) imagining Romeo and Juliet as Bobby and Betty (or *as played by* Bobby and Betty). This move by the defender of AAT is supposed to transfer the obvious perceptual nature of our perceiving Bobby and Betty to our imagining Romeo and Juliet, thus assuring the imagining to be genuinely perceptual, and thereby giving us real imagining rather than imagining by courtesy. Actors play the fictional characters, and this is supposed to enable us to imagine the fictional characters as (*played by*) the actors. But this does not touch the fundamental problem that abstract objects are just not subject to imagining, which is perceptual by nature. We cannot imagine Romeo and Juliet, *a fortior*i cannot imagine Romeo and Juliet as (*played by*) Bobby and Betty.

10 It is possible for the defender of AAT to resort to an externalist conception of intention and argue that such a conception requires the troop members to have some *de re rapport* with the fictional characters in order to intend to play them. This will bring up important issues concerning *de re rapport* but they are topics to be discussed on other occasions.

5 Imagination *de re*

The defender of AAT might challenge the legitimacy of the notion of imagination *de re* as applied to appreciation of narrative fiction. She might say that even though we often imaginatively engage ourselves when reading a narrative fictional story, such imaginative engagement should not be understood in terms of imagination *de re*, where the *res* is a fictional character, even if it is understood in terms of imagining a fictional character. She might say that in imaginatively engaging ourselves, we may be said to imagine Romeo kissing Juliet, among other things, but that imagining Romeo kissing Juliet is not imagining any particular entity kissing any particular entity, even though Romeo and Juliet are particular entities.

But this would be a peculiar view, to say the least. According to AAT, Romeo and Juliet are actually existing particular entities, and yet imagining Romeo and Juliet is not imagining particular entities.

We cannot imagine information but can imagine the carrier, or the encoder, of information (by imagining written words, for example). Likewise, it might be proposed, we cannot imagine Romeo and Juliet but can imagine the encoders of Romeo and Juliet by imagining Bobby and Betty, and this is good enough for our imaginative engagement.

Is imagining the encoders of Romeo and Juliet really a good enough surrogate for imagining Romeo and Juliet? No. The encoders of Romeo and Juliet are nothing more than presenters or representations of Romeo and Juliet, not Romeo and Juliet. The presenter or representation (or encoder) of information is certainly not the information itself. Imagining a piece of paper and the ink marks deposed on it is a far cry from imagining the conveyed information.

6 Other kinds of engagement

It is important to distinguish imagining Romeo from empathizing with Romeo. It is not just imaginative engagement with fictional characters that is important in our appreciation of narrative fiction. Empathetic engagement is equally important. But the latter is free from the problem of the former, even though there is a close connection between the two.

We can empathize with *x* only if we can put ourselves in *x*'s shoes, as it were. But we cannot put ourselves in *x*'s shoes unless we can imagine ourselves as being *x* in some appropriate sense of "imagine ourselves as being *x*". Given this, one might be tempted to think that putting ourselves in Romeo's (or Juliet's) shoes is

no more possible than imagining Romeo kissing Juliet. It would be a mistake to yield to this temptation. The crucial difference between the two kinds of engagement is a difference between the *de re* and the *de dicto*. Imagining Romeo kissing Juliet is *de re* imagining, where the *res* are Romeo and Juliet, whereas putting oneself in Romeo's (or Juliet's) shoes is *de dicto* in the sense of imagining ourselves to have the properties (and bear the relations) Romeo (or Juliet) has (and bears) in the story.

It is easy to confuse these two kinds of engagement and conclude mistakenly that since putting oneself in the fictional character's shoes (*de dicto*) is possible, imagining the fictional character (*de re*) is possible.

A related kind of engagement that may also be conflated with imagining fictional characters is that which is performed by thinking counterfactually: e.g., "If I were Romeo, I would kiss Juliet in that situation"; "If Romeo kissed me in that situation, I would slap him". Such counterfactual engagement is no more *de re* than empathy by way of putting oneself in the shoes of Romeo or Juliet. We are using the properties had by the fictional characters in the story.

There may probably be other kinds of engagement with fictional characters that might be conflated with imagining them and might be thought to be usable to defend AAT.[11]

7 Positive outlook

Our discussion has so far been negative, consisting of a sustained attack on AAT. It is time to look for something positive. Consider the following claims concerning imaginative engagement:

7. Our imagining Romeo kissing is imagining *de re*, where the *res* is Romeo.
8. Our imagining Romeo kissing is imagining an arbitrary entity that has exactly the properties Romeo has in the story, kissing.

7 and 8 appear plausible separately but incompatible with each other. AAT has Romeo as abstract, hence it makes 7 implausible, given the perceptual nature of

11 Suppose (i) that we can dream an abstract artifact to be kissing another abstract artifact. Suppose also (ii) that whatever we can dream we can imagine. Then it follows that we can imagine an abstract artifact kissing another abstract artifact. I am skeptical about (i) but admit that it is not easy to refute it. I, however, reject (ii). Dreaming allows much more than perceptually driven modes of content entertainment. This footnote was inspired by remarks by Kris McDaniel and Gonzalo Rodriguez-Pereyra.

imagining. This the objection to AAT we have been pursuing. But AAT could allow us to imagine Romeo in the sense of 8. 8 appears to abandon imagination *de re* in favor of imagination *de dicto*, so it appears to fail to honor an important function of fictional characters in our appreciation of narrative fiction. At the same time, 8 seems to be faithful to the phenomenology of our imagining Romeo kissing. This suggests a less negative attitude toward 8: We should try to simultaneously endorse 7 and save this phenomenology 8 highlights. That is, we should try to save both 7 and 8 in a coherent way.

One way to do so is to try to argue that an arbitrary entity having exactly the properties Romeo has in the story is Romeo. Of course, in order to do so, we need to do two things: (i) reject AAT and regard Romeo as being a kind of entity that can have – as opposed to hold, i.e., have in the story – those properties (being concrete, animate, human, male, etc.), and (ii) reject the somewhat intuitive view that Romeo is, if a human being, not just any old entity that has exactly the properties Romeo has in the story but one particular unique entity that has exactly those properties. The first of these two things is a natural thing to do for many who oppose AAT, especially for Meinongians (those who maintain that Romeo is a non-actually-existing human being). The second skirts the objection that Romeo is one, not many; for it implies that if two or more individuals had the properties in question, then none of them would be Romeo. Elaboration along the lines of (i) and (ii) is possible but unfortunately beyond the scope of this paper.[12]

Bibliography

Evans, G. R. (1982), *The Varieties of Reference*, Oxford: Oxford University Press.

Kripke, S. (1980), *Naming and Necessity*, Harvard: Harvard University Press.

Kripke, S. (2013), *Reference and Existence*: *The John Locke Lectures*, Oxford: Oxford University Press.

Sainsbury, R. M. (2009), *Fiction and Fictionalism*, Routledge.

Van Inwagen, P. (1977), "Creatures of Fiction", in *American Philosophical Quarterly*: 14(4), 299–308.

Van Inwagen, P. (1983), "Fiction and Metaphysics", in *Philosophy and Literature*: 7(1), 67–77.

Walton, K. (1978), "Fearing Fictions", in *Journal of Philosophy*: 75(1), 5–27.

Walton, K. (1990), *Mimesis as Make-Believe*, Harvard: Harvard University Press.

Walton, K. (1997), "Spelunking, Simulation, and Slime: On Being Moved By Fiction", in *Emotion and the Arts*, edited by M. Hjort and S. Laver, Oxford: Oxford University Press, 37–49.

12 For an attempt at such elaboration, see Yagisawa (2010), Chapter 6.

Yagisawa, T. (2010), *Worlds and Individuals, Possible and Otherwise*, Oxford: Oxford University Press.
Zalta, E. N. (1983), *Abstract Objects: An Introduction to Axiomatic Metaphysics*, D. Reidel.

Graham Priest

Objects That Are Not Objects

Abstract: When involved in projects concerning language and its limits, a number of philosophers (Wittgenstein, Heidegger, Frege) have been driven to the conclusion that there are certain things which appear to be both objects and not objects. They have tried to avoid the contradiction by various strategies, such as the apparently desperate one of declaring some of their own assertions to be meaningless. However, a quite different strategy is to accept the contradictions involved. This requires the use of a paraconsistent logic, though this is only a first step. How, given the resources of such a logic, can one understand the thought that something both is and is not an object? And does such a thing in any way destroy meaning? The paper explores these issues, and answers the questions.

1 Introduction

In the *Philosophical Investigations*, Wittgenstein says that:[1]

> the results of philosophy are the uncovering of ... bumps that the understanding has got by running its head against the limits of language

and adds, 'These bumps make us see the value of the discovery'. The point of this essay is to examine one of those bumps.

Many major philosophers have argued that there are things beyond the limits of language, things which language cannot talk about. Of course, there is a rub: if one argues that there are such things, one must, in the process, be talking about them. Contradiction stares one in the face.

One hardly needs to be a philosophical genius to see the contradiction involved; and the philosophers in question usually attempted some kind of philosophical evasive action. A notable one is the rather desperate strategy of the philosopher declaring his[2] own words meaningless. I think that, if one endorses the philosophical projects in question, there is a better solution: one should just accept the contradiction – or at least the contradiction which is generating the problem.

1 Anscombe (1968), p. 48.

2 Yes, all the ones I can think of were men.

https://doi.org/10.1515/9783110664812-012

In the first half of this essay, we will look at three prime examples of philosophers who, for closely related reasons, found themselves in this situation, and were inclined to this move. In the second part, we will look at the alternative. This presupposes a basic knowledge of paraconsistent logic, which many people will not have. So between these two parts, I insert an interlude explaining those basics.

2 The phenomenon

First, then, to the phenomenon in question. We will see three notable philosophers in whose thought this arises.

2.1 Wittgenstein

The first is the Wittgenstein of the *Tractatus*. In this, Wittgenstein puts forward an account of language, reality, and the relationship between them. The propositions of language are composed of names, put together with a certain form. Reality comprises states of affairs. These are composed of objects, put together with a certain form. And a proposition, p, describes a state of affairs, s, if the names in p refer to the objects in s, and the form of p is the same as that of s. (One might call this the *isomorphism* theory of representation.) However, the form of a proposition/state-of-affairs cannot be another object. If it were, a proposition/state-of-affairs would just be a congeries of names/objects. A form is the *way* that the names/objects are put together, something quite different.[3]

But now the rub. Form is not an object, and so cannot be the component of a proposition. One cannot, therefore, talk about form, which of course the *Tractatus* does at length. In his introduction to the English translation of the *Tractatus*, Russell comments:[4]

> Everything ... which is involved in the very idea of the expressiveness of language must remain incapable of being expressed in language, and is, therefore, inexpressible in a perfectly precise sense. ... [One may have] some hesitation in accepting Mr Wittgenstein's position, in spite of the very powerful arguments which he brings to its support. What causes hesitation is the fact that, after all, Mr Wittgenstein manages to say a good deal about what cannot be said...

3 For more on this and the following, see Priest (2002), ch. 12.

4 Pears and McGuiness (1974), p. *xxi*.

Wittgenstein was, of course, well aware of the issue, and it motivates the stunning conclusion of the *Tractatus*:[5]

> 6.54. My propositions are elucidatory in this way: he who understands me finally recognizes them as senseless, when he has climbed out through them, on them, over them. (He must so to speak throw away the ladder, after he has climbed up on it.) He must surmount these propositions; then he sees the world rightly.
> 7. Whereof one cannot speak, thereof one must be silent.

Wittgenstein gritted his teeth (albeit with evident pleasure) and says that claims of the *Tractatus* are meaningless.[6]

One cannot but admire Wittgenstein's chutzpah. But the move really will not wash. The claims of the *Tractatus* are not meaningless. You can read them and understand them; philosophers teach them to their students, who, likewise, understand them. To add insult to injury, the move saws off the very branch on which Wittgenstein is sitting. If the claims of the *Tractatus* are, indeed, meaningless, they cannot establish anything; in particular, then, they cannot establish that one cannot talk about form; and so they provide no reason for supposing that they are meaningless.

Here, then, is our first historical example.

2.2 Heidegger

Let us turn to the second: Heidegger. In the opening pages of *Being and Time*, Heidegger poses his *Seinsfrage*: what is *being*, what is it to be? This was the question that was to dominate his philosophical thinking, one way or another, for the rest of his life. But immediately after asking the *Seinsfrage*, he warns that, whatever being is, it is not, itself, a being. There is a difference of kind between being and beings: the so called ontological difference. He does not, there, give a reason for this, but it is a natural enough Neoplatonic thought. Being is the ground

5 Pears and McGuiness (1974), p. 74.

6 I note that the the full quotation from the *Investigations* in Section 1 above reads:

> the results of philosophy are the uncovering of one or another piece of plain nonsense and of bumps that the understanding has got by running its head against the limits of language.

So despite the obvious differences between the *Tractatus* and the *Investigations*, as far as the present matter goes, we do not seem to be in such different situations.

of beings; it is what makes them be; as such, it must be a quite different kind of thing.[7]

But now there is an obvious problem. To answer the question of being, one has to say something like: being is such and such. This puts being in a subject-position, and thus treats it as a being. So you can't answer the question of being. Heidegger came to accept this conclusion. An important line of thought in his later writing was that, though one cannot say what being is, art, poetry, and so on, can open people's eyes to being *showing* itself.

Be that as it may, matters are more desperate than this. It is not just that one cannot *answer* the question of being: one cannot even *ask* it. To ask 'what is being?' is itself to treat it as a being. Indeed, one cannot say *anything* about being. To say anything *about* it, would be to treat it as a being. Yet Heidegger's writings are replete with thoughts about being.

Heidegger, of course, realised that he had a problem, and essays various strategies to cope with it. One technique he attempted was "writing under erasure", saying it and crossing it out, as in:[8]

> ... a thoughtful glance ahead into the realm of *being* can only write it as ~~being~~. The crossed lines at first only repel, especially the almost ineradicable habit of conceiving *being* as something standing by itself ... Nothingness would have to be written, and that means thought of, just like ~~being~~.

The technique does not avoid the problem, though: even to explain what the crossing out means, Heidegger has to talk about being.

Evasion is, in fact, a forlorn cause. If being really isn't an object, then speaking about it is grammatically impossible, as Heidegger himself realises:[9]

> If we painstakingly attend to the language in which we articulate what the principle of reason [*Satz vom Grund*] says as a principle of being, then it becomes clear we speak of being in an odd manner that is, in truth, inadmissible. We say: being and ground/reason [*Grund*] 'are' the same. Being 'is' the abyss [*Abgrund*]. When we say something 'is' and 'is such and so', then that something is, in such an utterance, represented as a being. Only a being 'is'; the 'is' itself—being—'is' not. The wall in front of you and behind me is. It immediately shows itself to us as something present. But where is its 'is'? Where should we seek the presencing of the wall? Probably these questions already run awry.

7 For this and what follows, see Priest (2002), ch. 15.

8 Kluback and Wilde (1959), p. 81.

9 Lilly (1991), p. 51f.

Though not as forthright as Wittgenstein, the thought is essentially the same: Heidegger's statements about being are meaningless. And the problem with this is exactly the same as that for Wittgenstein. Heidegger's statements are not meaningless. We *do* understand them. Even when he explains what cannot be done, in the process doing it, we still understand him. Heidegger's words belie his claims.

This is the second example of our target phenomenon.

2.3 Frege

Let us turn to the third. This is to be found in Frege. Driven by the construction of his concept script (*Begriffsschrift*) Frege formulated an impressive and systematic philosophy of language. One of the problems he had to confront in doing this was what might be called *the unity of the proposition*.[10]

Consider a sentence like, *Socrates is sitting*. 'Socrates' refers to an object: the Ancient Greek philosopher we know and love. Objects are the referents of noun-phrases. 'Is sitting' refers to the property of sitting, or, as Frege calls it, the concept. Concepts are the referents of predicative phrases. However, *Socrates is sitting* is no mere list, <Socrates, sitting>; its two parts cooperate in some way to produce the unity expressed by the sentence.[11] How do they do this?

According to Frege, objects and concepts are quite different kinds of things. In particular, concepts are "unsaturated"; they have a "gap" in them; this can be filled by an object to create a unity. Thus, *Socrates* plugs the gap in the concept *is sitting* to deliver the appropriate unity.

But now we have a problem, and it is not that Frege's language of gaps and filling them is metaphorical: sometimes metaphors are all we have. Consider the concept *is sitting*. 'The concept *is sitting*' is a noun phrase, and so refers to an object, not a concept, as one might expect. This is awkward. For it means that the concept *is sitting* (that object) does *not* have a gap in it. What Frege really needs to say is something like: *is sitting has a gap in it*. The ungrammaticality of this is patent.

Frege is well aware of the point, and comments on it in a well known passage of 'Concept and Object' concerning the concept *horse*. He says:[12]

> I admit that there is a quite peculiar obstacle in the way of an understanding with my reader. By a kind of necessity of language, my expressions, taken literally, sometimes miss

10 For this and what follows, see Priest (2002), ch. 12.

11 One might naturally think of this as a proposition, though for Frege is was a truth value.

12 Geach and Black (1960), p. 54.

> my thoughts; I mention an object when what I intend is a concept. I fully realize that in such cases I was relying on the reader who would be ready to meet me half-way—who does not begrudge me a pinch of salt.

Frege is clearly embarrassed – but not embarrassed enough. The point appears to constitute a *reductio ad absurdum* of his whole theory. The claim that a sentence *Pa* expresses a unity because the object denoted by *a* fills a gap in the concept denoted by *P*, is just plain false. The concept denoted by *P* has no such gap.

Frege says no more about the problem. Wittgenstein and Heidegger offer a solution – albeit a Phyrric one. Frege has nothing to offer.

2.4 The underlying problem

Here then, are our three examples, three outstanding philosophers who hit the fact that, according to their own lights, they cannot really say what they want to say. And two of them, as least, explicitly or implicitly, declare their own words meaningless.

Prima facie, the situations of our three philosophers might appear to be quite distinct; but in fact, they are all manifestations of one and the same problem, namely that the special kind of thing that creates a unity out of objects, cannot itself be an object – but it is, because we refer to it in this way.

This is most obvious in the case of Frege. For him, concepts are the special kind of thing which create unities. They are not objects, or they could not do the job of unifying; but they must be, because we refer to them with noun phrases.

It is not difficult to see that Wittgenstein's problem is exactly of this kind. Form is the kind of thing which creates unities of names or of objects. As such it is not itself an object, but we refer to it as such. Indeed, Wittgenstein's problem is just the intellectual descendent of Frege's.[13]

It is not immediately obvious that Heidegger's issue is the same, but in fact it is. This is because he is working with an Aristotelian (pre-Fregean) notion of logical form. Statements are of the form: *S is* [*not*] *P*. There is, then, only one concept in Frege's sense, the copula, *is*, that is, *being*. The problem is a special case of Frege's.

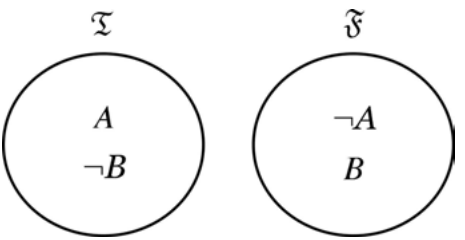
What we have seen, then, is that what is driving the predicaments of our three philosophers is the need to recognise something that is not an object – a principle of unity – as an object.

13 As noted, e.g., by Anscombe (1959), pp. 108 ff.

3 Interlude: Paraconsistent logic

So much for our target phenomenon. In the second main part of this essay I will turn to the question of how better one might respond to it. But first, as promised, a short primer on paraconsistent logic.

Let us start with so called classical propositional logic; that is, the logical theory invented by Frege, adopted by Russell, and then polished by a number of great logicians in the first half of the 20th Century. According to this theory, every situation – or interpretation as logicians call them – divides up all the sentence of the language in question into two: those that are true in the situation, $\mathfrak{T}$, and those that are false in it, $\mathfrak{F}$. Every sentence is in one or other of these, but not both. Negation, $\neg$, toggles a sentence between the two zones; so if anything is in one, its negation is in the other. We have, then, the following picture:



An inference is valid if in every situation in which the premises are in the $\mathfrak{T}$ zone, so is the conclusion. Said differently, it is valid if there is no situation in which the premises are in the $\mathfrak{T}$ zone, and the conclusion is not. So consider the inference of *Explosion*:

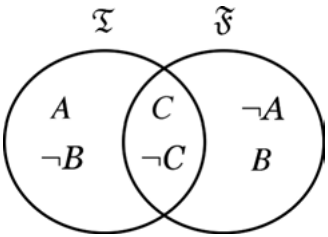
- $A, \neg A \vDash C$

This is valid, simply because there is no situation in which the premises are both in the $\mathfrak{T}$ zone; *a fortiori*, there is no situation in which the premises are both in the $\mathfrak{T}$ zone *and* the conclusion is not.

A paraconsistent logic is (by definition) one in which Explosion is not valid. There are many paraconsistent logics, but let me describe a simple one.[14] This is *exactly* the same as classical logic – with one modification. In a situation, the $\mathfrak{T}$ and $\mathfrak{F}$ zones may overlap. Negation still toggles a sentence between the two zones.

14 For a survey of paraconsistent logics, see Priest *et al.* (2015). The one described here is the logic *LP*.

If a sentence is in the overlap between the two zones, it is true and false, and so, then, is its negation. That is, it is in the overlap as well. Thus, we may have:



Explosion then fails to be valid. In the above diagram, both C and $\neg C$ are in the $\mathfrak{T}$ zone (and in the $\mathfrak{F}$ zone as well, but that is irrelevant), but B is not. So $C, \neg C \not\models B$. In virtue of the failure of Explosion, a paraconsistent logic can accommodate theories which are inconsistent, but in which not everything holds. Contradictions can, as it were, be isolated as singularities.

This is nearly all we need; there is just one more thing. In first-order logic, predicates have extensions and anti-extensions (in a situation). The things in the extension are things which satisfy the predicate; the things in the anti-extension are things which satisfy its negation. In the classical case, the extension and the anti-extension of a predicate are exclusive and exhaustive. But in the paraconsistent case, as one would expect, these may overlap.

One predicate, in particular, will concern us in what follows. This is the identity predicate, =. This is a binary predicate, and so its extension and anti-extension are sets of pairs of objects from the domain of objects, D. The extension of this predicate is as in classical logic, $\{\langle d, d\rangle : d \in D\}$. The anti-extension can be anything one likes, except that every pair, $\langle d_1, d_2\rangle$, must be in either the extension of the anti-extension. The extension of = is sufficient to deliver all the usual properties of identity. (The anti-extension is, as a matter of fact, completely inert in these matters.) However, the fact that the extension and anti-extension can overlap means that there can be situations in which things of the form $a = b$ and $a \neq b$ are both true.

4 Meaning and ineffability

With these preliminaries under our belt, let us now return to our target phenomenon. The problem of our three philosophers arose because of seemingly being forced to recognise some objects which are not objects. That is, of course,

a contradiction, but the resources of paraconsistent logic show how to make sense of it. Let us see how.

4.1 To be and not to be an object

First, what is it to be an object? Simply to be something; that is, simply to be in the domain of quantification. So we can define:[15]

- x is an object $:= \exists y\, y = x$

We may now reason as follows:

- $x = x$
- so $\exists y\, y = x$
- so $\forall x \exists y\, y = x$

That is, (unsurprisingly!) everything is an object.

But if x is not an object, we may also reason as follows:

- $\neg \exists y\, y = x$
- so $\forall y \neg y = x$,
- so $x \neq x$

What we see is that if x is an object that is not an object, $x = x$ and $x \neq x$.[16]

4.2 Naming

But does the fact that something is an object and not an object imply that claims about it are meaningless, or that one cannot talk about it? Let us investigate.

---

15 I do not intend the quantifiers here to be "existentially loaded". $\exists$ should simply be read a *some*. In Priest (2005) I use $\mathfrak{S}$ for the particular quantifier; and indeed, I think that it is better to write it like this. However, in the present context, this would merely distract from the matters at hand.

16 As a matter of fact, the conditional goes in the other direction as well. For suppose that $x \neq x$. For any y, either $y = x$ or $y \neq x$. In the second case, $y \neq x$ by the substitutivity of identicals. Hence, in either case, $y \neq x$. That is, $\forall y\, y \neq x$, and so $\neg \exists y\, y = x$: x is not an object.

Take some object that is not an object, and let 'n' be a name for it. The principle which naively (and correctly) governs the truth predicate, T, is the well-known T-Schema:

- $T\langle A\rangle \leftrightarrow A$

The angle-brackets here are a name-forming device. Of course, many who have addressed the semantic paradoxes of self-reference have taken the Schema to have restricted validity. I do not. However, this is not the place to go into that matter.[17]

Now, other semantic notions are governed by similar schemas. In particular, if D is the denotation relation, it is governed by the D-Schema:[18]

- $\forall x(D(\langle t\rangle , x) \leftrightarrow t = x)$

Here, t is any term, and $\langle t\rangle$ is its name.

Instantiating the D-Schema, we get:

- $D(\langle n\rangle , n) \leftrightarrow n = n.$

And since $n = n$, it follows that $D(\langle n\rangle , n)$. That is, as one would expect, $\langle n\rangle$ is a name for n. However, since n is not an object, $\neg\exists y\, y = n$. That is, $\forall y\, y \neq n$, and so for any term, t, $t \neq n$. Again instantiating the D-Schema again we get:

- $D(\langle t\rangle , n) \leftrightarrow t = n.$

And assuming that this biconditional contraposes, we have:

- $\neg D(\langle t\rangle , n)$

One might have worries about the contraposibility of the biconditional here: there are certainly plausible reasons for supposing that the biconditional of the cousin of the D-Schema, the T-Schema, does not contrapose.[19] But let is set those issues aside here. What the last displayed statement shows us is that no term, t, is the name of n. n has not name – not even 'n'!

17 I have defended the T-Schema in many places. For a start, Priest (2006b).

18 See, e.g.,Priest (2005), ch. 8.

19 See, e.g., Priest (2006a), 4.9.

4.3 Statements about *n*

What follows from this? First, it does not follow that statements about *n* are meaningless. They are meaningful. Even if *n* is not an object, '*n*' is a perfectly good syntactic name, and so statements deploying it – such as *Pn*, where *P* is some monadic predicate – are are perfectly grammatical. Not only are such statements grammatical, their meaning is perfectly clear, and is given by the standard truth conditions. Thus, *Pn* means that *n* satisfies '*P*'.

So Wittgenstein and Heidegger were wrong about this matter (of course!). One can talk perfectly meaningfully about objects that are not objects.

However, that is not an end of the matter. For something of the form *Pm* to be about an object, '*m*' must be a name for it. But we have seen (assuming the contraposibility of the *D*-Schema) that *n* has no name. So one can say nothing about *n*. Statements about *n* are ineffable: one can say nothing about *n* – even though one can! So Wittgenstein and Heidegger were right about this matter, the objects they wished to talk about (and did) were ineffable.

What we have seen, then, is that if our three philosophers had simply accepted that the things they were dealing with were objects that were not objects, they could have had their cake and eaten it too. They would indeed have been trespassing into the ineffable.[20] But there is no reason to suppose that the things said are meaningless; nor any temptation to the self-destructive thought that they are so.[21]

Of course, some will see being a dialetheist, and accepting contradictions – such as that something is and is not an object – as an unacceptable cost. This is not the place to go into that matter.[22] So let me just put it this way. You come out of a lecture by Wittgenstein or Heidegger in 1925. Which is worse to have to say? – 'Well, you know, he actually contradicted himself', or 'Well, his lecture was totally meaningless'?

4.4 König's paradox

We have been dealing with some hard-core metaphysics, and it might well be thought that the kind of knots we have been exploring are restricted to such

20 Further on this theme, see Priest (2002).

21 Interestingly, there is evidence showing that in his later years, and mostly in his private diaries, Heidegger actually because a dialetheist in the matter at hand. See Casati (2016).

22 On which, see, for example, Priest (2006b).

matters. Let me conclude by pointing out that they are not. They occur just as much in logic and set theory.

There are many paradoxes of self-reference, but let us consider one of them: König's Paradox.[23] This concerns ordinals. Ordinals are numbers that extend the natural numbers, 0, 1, 2, … into the transfinite. Thus, after all the natural numbers, there is a least infinite ordinal, ω, and then a next $\omega + 1$, and so on. So the ordinals look something like this:

- $0, 1, 2, \ldots \omega, \omega + 1, \omega + 2, \ldots 2\omega, 2\omega + 1, \ldots \omega^2, \ldots$

A crucial feature of the ordinals is that they preserve the property of the natural numbers that any collection has a least member. This is called the *well-ordering property*.

Now, how far the ordinals go on, is a vexed question, both mathematically and philosophically. However, it is beyond dispute that given any language with finite basic resources, such as English as it is here and now, there are ordinals that cannot be referred to by a non-indexical noun-phrase of that language. This can be proved by a perfectly rigorous mathematical argument concerning cardinalities. Hence, there are ordinals that cannot be referred to in this way; and so by well-ordering property, there is a least such ordinal. Call this α (or let α be an abbreviation for the description 'the least ordinal that cannot be referred to'). Then since one cannot refer to α, one can say nothing about it—it is ineffable. But one can say things about it – for example that it is the least ordinal that cannot be referred to.

This is König's Paradox. I am not suggesting that it is driven by the considerations of unity that drive the conundrums of our three target philosophers. It is not. I give it merely to show that the strange territory that our three philosophers would be lead into if they became dialetheists is already to be found in set theory and logic.

5 Conclusion

We have been exploring the unfamiliar world of objects that are not objects, and we have seen how the thoughts of three of the most important philosophers of the last 150 years take us into it. I have not tried to defend the views of those philosophers – or the parts of them which lead in this direction. Whether or not one

23 See, e.g., Priest (2002), pp. 131 ff.

should do so is a matter for an entirely different occasion.[24] This point of this paper is simply to show that dialetheism opens up a whole new world of possibilities – or, one might quip, impossibilities – concerning profound areas of metaphysics.[25]

Bibliography

Anscombe, G. E. (1959), *An Introduction to Wittgenstein's Tractatus*, London: Hutchinson and Co.

Anscombe, G. E. (tr.) (1968), *Philosophical Investigations*, Oxford: Basil Blackwell.

Casati, F. (2016), *Being. A Dialetheic Interpretation of the Late Heidegger*, PhD thesis, University of St Andrews.

Geach, P.and Black, M. (trs.) (1960), *Translations from the Philosophical Writings of Gottlob Frege*, 2nd edn, Oxford Basil Blackwell.

Kluback, W. and Wilde, J. (trs.) (1959), *The Question of Being*, Harrisonburg, VA: Vision.

Lilly, R. (tr.) (1991), *The Principle of Reason*, Bloomington, IN: Indiana University Press.

Pears, D. and McGuiness, B. (trs.) (1974), *Tractatus Logico-Philosophicus*, London: Routledge and Kegan Paul.

Priest, G. (2002), *Beyond the Limits of Thought*, 2nd edn, Oxford: Oxford University Press.

Priest, G. (2005), *Towards Non-Being*, Oxford: Oxford University Press, 2nd edn, 2016.

Priest, G. (2006a), *In Contradiction*, 2nd edn, Oxford: Oxford University Press.

Priest, G. (2006b), *Doubt Truth to be a Liar*, Oxford: Oxford University Press.

Priest, G. (2014), *One*, Oxford: Oxford University Press.

Priest, G., Tanaka, K. and Weber, Z. (2013), "Paraconsistent Logic", in *Stanford Encyclopedia of Philosophy*, edited by E. Zalta, http://plato.stanford.edu/entries/logic-paraconsistent/.

24 There is some discussion of the matter in *One* (Priest (2014)), which gives my own reasons for supposing that there are objects that are not objects. I note that even though the present essay was written after that book, it actually provides a bridge between *Beyond the Limits of Thought* (Priest (2002)) and *One*.

25 This paper is a written up version of a talk given in a number of places including: the Korean Society for Analytic Philosophy, the Jowett Society (Oxford), Harvard University, Syracuse University, and the CUNY Graduate Center. I am grateful to the audiences for their thoughtful and helpful comments.

Part IV: **Composition, Organisms, and Persons**

Peter Simons

The Concept of Organism and Degrees of Composition

Abstract: The idea of composition – one thing's being made up of several things – is age-old, but we owe focus on the question *when* things compose other things to Peter van Inwagen. Philosophers' answers to this, his Special Composition Question, have been very varied. There are two extreme positions: nihilism and universalism. Nihilism says that nothing is ever composed of anything else, while universalism says that any things compose another thing. Both answers have been defended, and both fly in the face of common sense and much evidence from science. Intermediate positions however prove to be difficult to formulate and defend. Peter's own intermediate position is towards the nihilistic end of the spectrum. According to it, the only cases of composition are when simples compose a living organism. While sharing many of the difficulties of nihilism, it faces its own special composition difficulty: what it is to be a single organism is vague. I adduce evidence from biology that what biologists (unhappily) call "organismicity" or "organismality" comes in degrees. It is clearest in the case of macroscopic animals, but becomes much murkier among the other biological kingdoms. Further, nearly all cases of major evolutionary advance have occurred through the coming together of previously independent individuals, from replicators to eukaryotic cells to multicellular organisms. Transitional cases therefore likewise stand in the way of a clear-cut criterion of composition. I do not shrink from this vagueness – I welcome it, for it matches the evidence from other domains than the organic, that composition is not an all-or-nothing affair, and that what it is to count as a single, complex individual is not an exact matter. This is not a weakness of our conceptions or theories, but a feature of the world, and the sooner we work with rather than against it, the quicker our metaphysics of composition will recover and become relevant to real-world cases.

1 The special composition question and some extreme answers to it

When do several objects compose another object? This simple question was put in all clarity in the book *Material Beings* (hereafter: *MB*) by Peter van Inwagen (hereafter: *PvI*), where he called it the Special Composition Question (hereafter: *SCQ*). Answers there are several. Two extreme answers are: "Never" and "Always". The former answer, compositional nihilism, is unpopular, but there are and have been those who believe it. The latter answer, compositional universalism, is much more popular among metaphysicians, and was indeed the position adopted by

https://doi.org/10.1515/9783110664812-013

the father of mereology, Stanisław Leśniewski. Both extreme answers have their severe difficulties, and I do not intend to spend time arguing against them, but shall assume, with *PvI*, that a more moderate answer, somewhere in between, is correct. His own answer is indeed in between, but hardly moderate, since he considers the only things that are composite are organisms, individuals having a life. This is compositional organicism. According to organicism, while there are animals, plants, bacteria and other living things, they do not strictly have such intermediate parts as legs, lung or livers, cell nuclei or cell membranes. Away from organisms, there are no stars, planets, continents, mountains, rocks, cities, buildings, trains, tables, books, coins, molecules or atoms. All such inorganic composites and non-organismal parts of organisms are rejected in favour of there being collections of mereologically simple objects arranged in this or that way.

I make no secret of disagreeing with *PvI* on this and regard all those kinds of things just listed as genuine composite individuals. However, I do not intend to argue at length for their existence in this paper. Rather I want to concentrate on those composite things that *PvI* thinks *do* exist, namely organisms, and raise a number of issues about them. The intention then is (slightly) more irenic: it is to attempt to assist us in thinking about what organisms are and when they exist, and consider what implications this has for metaphysics generally and mereology in particular. I shall here and there betray my non-organicist view by talking about parts about whose existence *PvI* would disagree.

When I started this paper I fancied myself apprised of all varieties of mereological folly, but recently came across another answer which is so far off the wall that I never even considered it. It is what my Trinity colleague Kenneth Pearce calls 'mereological idealism', which he formulates as: "co-apprehension of the constituent parts is a metaphysically necessary condition for the existence of a composite object" (Pearce (2017), p. 203). The originator of this view, is, appropriately enough for Trinity idealism, George Berkeley, who wrote in *Siris* (1744), "it is the mind that maketh each thing to be one". To refute this amazing view is matter for another time: I mention it only to demonstrate that there are not two but three views which are manifestly more extreme than *PvI*'s organicism.

I shall be introducing a number of theoretical but still empirical considerations about organisms which originate not with mereologists or metaphysicians but with biologists themselves. The aim will be to argue in the light of these considerations that organicism could by no means be a straightforward position, and that there are legitimate reasons to think that, if organicism were to be maintained, it would have to be a much more nuanced and differentiated account than we might initially think. Ultimately though, I consider that the aporiai thereby raised serve to throw serious doubt on the tenability of organicism as a doctrine.

2 A few biological facts

Biologists themselves are unclear and uncertain, and disagree about what it is to be an organism, so before delving into their areas of uncertainty and disagreement I want to start with a few undisputed facts. Firstly, life on Earth is divided into two superkinds: the prokaryotes and the eukaryotes. Prokaryotes are small single-celled organisms with a simple cell structure, lacking the nucleus and membrane-bound organelles of eukaryotes. They are 10–20 times less in diameter than eukaryotes and a typical genome is about one thousandth as big as that of a eukaryote. There are two subdivisions of prokaryotes, bacteria and archaea, which differ in genetics and chemistry. They typically reproduce, by binary fission, in around 1–3 hours. Although small, prokaryotes are extremely varied and versatile and occupy many more environmental niches than eukaryotes, which they outweigh in biomass by a factor of around ten. A single gram of soil typically contains around 10^{10} (ten billion) prokaryotes belonging to some 8 million species. They were the first life to evolve between 3 and 3.5 billion years ago (bya) and will probably outlast everything else on the planet. Extraterrestrial life, when we find it, will almost certainly be predominantly prokaryotic in kind, if not in all chemical detail like ours. Prokaryotes are also indispensable to the functioning of many eukaryotes, for example by enabling digestion. Eukaryotes evolved around 1–1.5 billion years later and comprise everything else: protists, fungi, many algae, plants and animals. They typically reproduce not by fission but by meiosis and mytosis, allowing DNA recombination.

It is now generally accepted that eukaryotes evolved through endosymbiosis, that the mitochondria and chloroplasts of eukaryotes were originally free bacteria whose invasion of probably archaean prokaryotes gave both host and parasite evolutionary advantages. The mitochondria of animals and chloroplasts of algae and plants are genetically distinct from the eukaryote's own genetic material contained in cell nuclei.

Another division among organisms is that between unicellular and multicellular. With a few exceptions, prokaryotes are unicellular and live independently, only a few (such as *Escherichia coli*) forming colonies with some mild specialisation. Some small animals, the protists (such as amoebae), are unicellular. The most familiar multicellular organisms are animals, plants and fungi. An adult human body is estimated to be composed of around 37 trillion cells, though the number of microorganisms of other species inhabiting that body is thought to be between 1 and 4 times as great. Multicellular eukaryotes are the most recent main form of life to appear, but it is considered that multicellularity has evolved independently several times.

While single prokaryotes and multicellular animals are the most familiar unified organisms, it is important to observe that there are forms of life which vary to a greater or lesser extent from this paradigm. Colonial organisms such as corals, sponges and syphonophores, the most famous of which is the Portuguese Man o' War (*Physalia physalis*), consist of numerous cospecific animals (zooids) which may live separately but form large unified bodies and in many cases exhibit functional specialisation as well as apparently purposive behaviour, despite lacking a nervous system or organs. There are many symbiotic relationships in which organisms of different species exist together for their mutual benefit and in some cases cannot exist or reproduce separately. Lichens are a famous case, consisting of algae or cyanobacteria living among the filaments of various fungi in a symbiotic relationship. It is thought that around 6% of land is covered by lichens. Recall that eukaryotes probably evolved through symbiosis. Eusocial insects such as honeybees, ants and termites form colonies of genetically identical but caste-specialised individuals whose collective behaviour is sufficiently similar to that of standard organisms that biologists have coined the term 'superorganism' for them. Another form of aggregation is biofilms, in which multiple species of microorganisms bind together to a surface and are held by a chemically secreted matrix. The most familiar sort of biofilm is the one your dental hygienist will tax you about: the dental plaque that can cause damage to teeth and gums. Adding the third dimension, microbial mats are aggregations of microorganisms several layers deep, and were probably the first forms of life to leave fossils, around 3.5 bya, and the first colonists of land around 1.2 bya.

What we instinctively think is an individual organism sometimes is not. A mushroom or toadstool is not an organism, but the fruiting body of a fungus, much of which exists below the soil, and is connected by mycelia or threads. The parts are genetically identical and there is communication between them, so the larger self-cloning individual is organism-like. In 2003 Catherine Parks of the US Forestry Service in Oregon and colleagues published an account of a specimen of the dark honey fungus *Armillaria ostoyae* (post-2008 *A. solidipes*) which covers 965 hectares or 10 sq km and is counted as the largest single living thing found to date. It is estimated to be between 2,400 and 8,600 years old. It regularly produces (borderline) edible mushrooms every autumn, all over the forest. Plant clones also throw up problems. Dandelions reproduce by cloning: plants such as infest my lawn are counted by some biologsts not as individual organisms but as distributed parts of a single genetic individual. In a mountainside in Fishlake National Forest, Utah, there is a grove of some 47,000 male quaking aspen trees (*Populus tremuloides*) covering 43 hectares, weighing over 6,000 tons, genetically identical, sharing a single root system and estimated to be over 80,000 years old. It is known as 'Pando'. The trees, like mushrooms, are visible individuals, but the biological

individual appears to be the whole grove. Biologicists have coined the term 'genet' to denote such clonal colonies of genetically identical units.

Finally, I must mention the heterogeneous polyphyletic group of slime moulds, which exhibit weirdly alien characteristics. Plasmodial slime moulds consist of many nuclei inhabiting a common structure of cytoplasm without interior cell walls. By contrast, cellular slime moulds consist of amoebic protists which can and typically do live separately, but when food is scarce they respond to a molecular secretion and congregate together into a unified grex, a slug-like body which functions like a single organism, moving to find food and able to produce a spore-producing fruiting body like a fungus, during which approximately 20% of the amoebae die. The most famous slime mould is *Dictyostelium discoideum*, which feeds on *E. coli* and whose genome has been completely sequenced.

The points of this empirical detail are threefold. The first is simply to recall the wide diversity of forms of life. For obvious reasons, non-biologists tend to think first of those organisms that are most familiar, but away from these paradigms, life goes on in many and often strange ways. The second is to provide data for the following section, in which biologists' concerns about the definition of 'organism' are discussed. The third and metaphysical point is to provide us with a range of examples against which to test the Inwagenian organicist hypothesis.

3 Biologists worrying about the organism concept

The variety of forms of life that we have sketched is one factor in causing something of a minor crisis in theoretical biology, which has been going on for some years, and seems not to have resulted in a consensual outcome. While the majority of biologists are agreed that there are paradigm cases of single, unified organisms; multicellular animals being one, and single prokaryotes at the other end of the scale, it is far from clear that a single set of necessary and sufficient conditions can be drawn up which capture all and only single organisms in their extension. To this end, biologists have coined the two ugly terms 'organismicity' and 'organismality' as placeholders within which characteristics of being an organism can be discussed with minimal prejudice to a theoretical decision (or lack of one). (I shall use the more recent and slightly less ugly 'organismality'.) One of the results of this has been to suggest that organismality is not an all-or-nothing affair, but that there may be different degrees of it, and that there may in fact be more than one dimension of variation across which complex entities exhibit degrees of organismality. Indeed some writers consider the notion of organism to be otiose in biology. Wilson (2000) argued that "Biology lacks a central concept

that unambiguously marks the distinction between organism and non-organism because the most important questions about organisms do not depend on this concept", while Margulis and Sagan contend that "the completely self-contained individual is a myth that needs to be replaced with a more flexible description." (Margulis and Sagan (2002), p. 19). At all events, despite the standard view being that, to quote Goodwin and Dawkins), "Biology is the study of life, and life comes in the form of organisms. One might expect, then, to find in biology some generally agreed description of what an organism is" (Goodwin and Dawkins (1995), p. 47) – one does not. The quote comes from a paper with the significant title, "What is an organism?: a discussion." A glance at the titles of other papers than these also shows the concern (cf. e.g. Pepper and Herron (2008); Folse and Roughgarden (2010)). Individual cases also throw up doubts (cf. e.g. Jantzen (1977); Anderson (2000)).

It is not that biologists are trying to do philosophy or merely splitting terminological hairs. Contrast this debate with the more terminological question as to whether viruses are organisms, which is usually shrugged off with comments like "in some ways they are like organisms, and other ways they are not". Biologists are looking for an operational definition, a way or ways to determine empirically whether a specimen in the lab or the field is an organism, or a part of an organism, or a collection of organisms. None mention viruses in this context.

Since these biologists discuss a wide range of examples (such as those I gave above) but also differ considerably in the answers they offer as to what does or does not constitute organismality, I shall concentrate on just one proposed answer, which has been put forward by David Queller and Joan Strassmann of Washington University, St. Louis. Their papers attract large numbers of citations so they may be considered central in the discussion. Both work in specialised areas such as symbiosis and conflict, so their interests are linked to the theoretical question. In their paper "Beyond society: the evolution of organismality", (Queller and Strassmann (2009)), they write (Abstract):

> The evolution of organismality is a social process. All organisms originated from groups of simpler units that now show high cooperation among the parts and are nearly free of conflicts. We suggest that this near-unanimous cooperation be taken as the defining trait of organisms. Consistency then requires that we accept some unconventional organisms, including some social insect colonies, some microbial groups and viruses, a few sexual partnerships and a number of mutualistic associations. Whether we call these organisms or not, a major task is to explain such cooperative entities, and our survey suggests that many of the traits commonly used to define organisms are not essential. These non-essential traits include physical contiguity, indivisibility, clonality or high relatedness, development from a single cell, short-term and long-term genetic cotransmission, germ–soma separation and membership in the same species.

The degree of cooperation and the degree of conflict among interacting individuals (of whatever kind or kinds) allow a two-dimensional matrix of continuous variation to be set up for different cases, where it can be groups of cells, individuals in groups of multicellular individuals, and individuals of different species that are cooperating or in conflict.

I do not endorse their theory fully, especially such rather marginal views as that a lifelong monogamous pair of albatrosses (that may spend much of the year thousands of kilometres apart) constitute something we could even consider calling a single organism. However, it does seem plausible to suppose that increased cooperation and reduced conflict among once separable and independent entities will tend to be conducive, since it often confers evolutionary advantage, to fostering transitions from collections of individuals to more complex single organisms, with specialisation and symbiotic mutual dependence causing loss of self-sufficiency among former symbionts. The assemblage of replicators in prokaryotes, of prokaryotes in eukaryotic cells, and of eukaryotic cells in multicellular eukaryotes, are, as Maynard-Smith and Szathmáry put it in the title of their important book, *The Major Transitions in Evolution* (Maynard-Smith and Szathmáry (1995)).

The take-home point here is not whether Queller and Strassmann are correct in all respects. For example they are prepared to consider organismality combining more than one species, whereas some authors consider genetic homogeneity another significant factor. The point is rather is that cooperative, competitive and conflicting behaviours in and between organisms and their parts all come in degrees, and the crucial evolutionary transitions mentioned above cannot have been instantaneous. So there are degrees of organismality lying strictly between all and nothing, to which such more or less exotic cases as honeybee colonies, slime moulds, biofilms, aspen groves and syphonophores offer contemporary witness.

4 Implications for the special composition question

Let me start by recalling *PvI*'s answer to the *SCQ*. It is that

> For some y, the xs compose y iff the activity of the xs constitutes a life

(*MB*, p. 82. I have left out the case where there is only one x. A single simple could not have a life – unless one were to agree with Leibniz on the existence of monads.)

By 'life' *PvI* means the biological life of an individual organism. For the moment I shall simply explore what the cases and disputes I have mentioned may mean for this position. To make it easy to refer back, I shall number the theses I intend to uphold.

1. All organisms have parts which are not organisms.

   PvI agrees with this, because his simples do not have lives. Assuming that everything that has parts is ultimately composed of simples, I agree with his point.

2. Some organisms have other organisms as parts.

   PvI seems to agree with this. I consider it true of all multicellular organisms, since their constituent cells have lives of their own: they are engendered, live for a while, and die. Call such organisms which are proper parts of others *mereo-organisms*. They do not include such organisms as inhabit a larger one without being parts of it, gut microbiome or parasites for example.

3. All organisms have microscopic parts which are neither simples nor organisms.

   PvI denies this, while I affirm it: protein molecules and DNA molecules are the kinds of parts I have in mind.

4. All organisms are partly constituted by stuff (not individuals) which are not organismic.

   Water, protein, cytoplasm are universal in terrestrial organisms, while cellulose and blood are found in some but not all. *PvI* disagrees.

5. Some organisms have macroscopic parts which are neither simples nor organisms.

   I count such things as leaves, roots, bones, feathers, scales, teeth, limbs, brains and livers in this category. *PvI* must disagree. I have yet to summon up the courage to tell my dentist that, according to a respected philosopher colleague, strictly and philosophically speaking, no one has any teeth.

6. Some organisms form aggregates or ensembles which are not organisms.

   Biofilms and early embryos are examples. *PvI* can agree with this and clearly does with the second kind.

7. Some organisms arise or arose from aggregates or ensembles that were not of previous organisms.

   I take it this applies to the earliest forms of life, which arose by the combination of replicators, proteins and whatever (no-one knows the story). *PvI* agrees with this.

8. Some organisms are members of ensembles which are similar to organisms (exhibit high organismality).

   Lichens, slime moulds, eusocial insects and siphonophores are examples.

9. Some organisms have macroscopic parts with high organismality.

   Aspen groves and giant fungi are examples.

10. Some organisms have microscopic parts with moderate to high organismality.

    The mitochondria and chloroplasts of eukaryotes are examples.

11. It is vague as to when organisms come to be and cease to be.

    We both think this. Generation and death are never instantaneous (though the latter can be quick).

12. It is vague what microscopic parts an organism has at any time at which it is alive.

    Respiration, metabolism, growth, digestion, decay go on in organisms all the time. For *PvI* the vagueness concerns only simples and mereo-organisms, if there are any. For me it's wider.

13. Composition is vague.

    That is, it is vague in many cases as to whether a certain multitude of objects compose another object at a given time. This follows from 11 and 12, and we both agree that it is true.

14. Composition comes in degrees.

    I think this is a natural thing to infer from 11, but also I think from 7–10. I also think, unlike *PvI*, that it applies in the inorganic sphere, to wholes such as stars, planets, houses and motorcycles.

15. Identity is vague.

    PvI affirms this in Chapter 18 of *MB*. I disagree with this: identity does not for me come in degrees (this must be argued elsewhere). More surprisingly perhaps, I am happier with

16. Existence is vague.

    PvI affirms this in Chapter 19 of *MB*. This seems to follow from the following: if it is vague whether the *x*s compose an individual *y* at time *t*, then it's vague whether *y* exists at *t*, providing of course no other collection actually does compose *y* at *t* (so the *x*s are, as it were, the prime candidates for the composing). Suppose the *x*s go some way towards composing *y* such that there are times at which it's vague whether *y* exists at them, but the *x*s then retreat in some way from composing *y*. Then it's vague whether *y* existed at any time, so it's vague whether *y* existed (period). The existence of the *x*s is not in dispute, nor is it ever other than false that the *x*s are identical with some individual *z* (they are many, not one).

Many of these positions are irenic in relation to *PvI*'s organicism, so while I continue to disagree with that position, the points of contact and agreement are several and important. Another respect in which we are in agreement is in the proposition that there is in general a causal element in parthood (*MB*, p. 81). I shall not expand on that, but we agree that it rules out not only compositional universalism but some of the more simple-minded suggestions for when composition occurs, such as spatiotemporal proximity. Aristotle says in the *Metaphysics* that a bundle of sticks is one because it is tied together (1016a1–2). That is an extremely weak form of unity. But what makes one of the sticks one is that it is moderately rigid (not continuous, as Aristotle assumes): its parts cohere so that it moves all together under not too extreme circumstances. That cohesion, which also applies within organisms, is causal, and turns on the exchange of photons holding molecules together and of gluons holding baryons together. While the operation of certain forms of causality (such as the strong nuclear force) requires spatial proximity of the interacting items, it is not because they are close together that they may compose a whole. In this then I agree with the negative parts of *MB*. Rather, it is the causality that binds, and that kind of causality requires spatial proximity. Other instances of causality may not. As to what makes a baryon a mereological simple but with several properties, that is another story (Simons (2005)).

Despite the several points of agreement however, the point of my highlighting the widely felt (among biologists) unclarity around the concept of an individual organism remains to give scientific reasons, ones that impinge internally on the position itself, to suppose that

17. Organicism is for empirical reasons an unstable and untenable answer to the Special Composition Question

and that a more liberal yet still intermediate position is to be preferred. The next section gives additional reasons to hold this view.

5 Wherefore intermediate wholes?

Throughout the discussion of organismality, I have differed from *PvI* in affirming the existence of objects that are mereologically intermediate between organisms at the top and simples at the bottom: atoms, molecules, cell parts, tissues, organs and so on. *PvI* will reply that he can cover what I and like-minded intermediarists are committed to by the general recipe of 'simples arranged *F*-wise', for suitable choices of the term or predicate *F*. I have two related concerns about this. One is that scientists engaged in the study of intermediate wholes in any particular field employ them or talk about them because they play important roles in scientific explanations. Cell membranes selectively filter chemicals into and out of cells, tissues hold macroscopic parts of organisms together, afford them rigidity, promote locomotion, protect them from infection etc., the chitin exoskeletons of crustaceans afford them protection and lend them corporeal support and rigidity, the leaves of trees promote respiration and photosyntesis. They are parts with specific functions. Secondly, in the inorganic sphere, intermediate parts are essential tokens in engineering design: the parts of a mobile artefact such as a car or aero-engine are designed and manufactured precisely to contribute via their properties and the ways in which they need to fit and operate together to the functioning of the whole. In advanced engineering, the science and engineering work together. A good example is afforded by the turbine blades of Rolls-Royce aero engines such as the Trent XWB turbofan, which despite their size and geometric complexity are single crystals of a nickel–aluminium alloy able to withstand temperatures of over 1700°C, higher than the melting point of nickel itself, and angular speeds of more than 12,000 rpm. They can do this in part because of the molecular characteristics of the alloy, in part because they are honeycombed with channels affording cooling, and in part because they are coated with a heat-resistant ceramic material. In designing such components, engineers employ knowledge about the properties of parts and features (molecules, crystals, cooling channels, grooves etc.) intermediate in scale between the whole engine and its ultimate constituent simples.

These points cut no metaphysical ice for *PvI*, since he does not insist that we layfolk or other experts talk the metaphysically basic and apt talk in our lives, but

for my part I consider that if scientists and engineers make seemingly indispensable use of intermediate wholes in their functional explanations and designs, I am not prepared for the sake of a speculative thesis to claim they are chasing metaphysical mirages. I certainly do not have the nerve to go to Rolls-Royce at Derby and tell the assembled engineers that, according to a respected colleague, yea, even a metaphysician, none of the things they fondly think they deal with strictly and philosophically exist. Their response would be too predictable: So much the worse for metaphysics. But I love metaphysics, and wish it to prosper. To do that, it must go out to meet theorists and practitioners rather than downgrade or snub their conceptions.

6 The process aspect

The simples and organisms of organicism are continuants, that is, objects enduring through time. Their interactions, operation, coming to be and passing away are however all occurrents, that is, events and processes which perdure through time, i.e. persist through the accretion and continuation of new temporal parts, in all cases (in my view) causally driven. Some philosophers of biology are process ontologists: the most prominent example being John Dupré (*vide* Dupré (2012)). According to Dupré, we organisms are all processes. I disagree (Simons (2018)). We survive and persist in a different way from processes, though we are clearly closely dependent on them (that is what physiology is about). Indeed, in my view, all continuants are ontologically dependent on and posterior to processes which constitute (but do not compose) them (Simons (2000)). The relation between temporal parts of such processes as constitute continuants I call, following Kurt Lewin, *genidentity*. Genidentity is a form of stable causal continuity among process phases.

Now, suppose I am right that continuants, including organisms but also quarks and leptons etc., are dependent on and constituted by underlying processes, what story should we tell about composition as it applies to processes, in particular those which make up the life of an organic individual? Clearly, this is not a simple affair, as a glance at physiology and biochemistry will reveal. Nor is it one on which philosophers are by trade qualified to pronounce. However, one very general connection does fall within ontology. If a continuant P is part of a more comprehensive continuant W at time t, then this must be because the processes constituting P at t are part of or are among the processes constituting W at t. That means that we have to work harder at the mereology and combinatorics of processes. This is hard, because we are ingrainedly fixated on continuants. Process mereology has barely been discussed, and I do not claim to be much wiser

than anyone else about how it might go. Clearly there will be a strong causal element to any sensible, intermediate account of when and how processes are parts of or compose others. All I can and will say here and now is that this is a subject for another time, and for much more work.

Bibliography

Anderson, J. O. (2000), "Evolutionary Genomics: Is Buchnera a Bacterium or an Organelle?", in *Current Biology*: 10, 866–868.

Dupré, J. (2012), *Processes of Life*: *Essays in the Philosophy of Biology*, Oxford: Oxford University Press.

Folse, H. J. and Roughgarden, J. (2010), "What is an Individual Organism? A Multilevel Selection Perspective", in *The Quarterly Review of Biology*: 85, 447–472.

Goodwin, B. and Dawkins, R. (1995), "What is an Organism?: a Discussion", in *Perspectives in Ethology*, Vol. 11, *Behavioral Design*, edited by N. P. Thompson, New York: Plenum Press, 47–60.

Jantzen, D. H. (1977), "What are Dandelions and Aphids?", in *The American Naturalist*: 111, 586–589.

Margulis, L. and Sagan, D. (2002), *Acquiring Genomes. A Theory of the Origin of Species*, New York: Perseus Books.

Maynard-Smith, J. and Szathmáry, E. (1995), *The Major Transitions in Evolution*, Oxford: Freeman.

Pearce, K. L. (2017), "Mereological Idealism", in *Idealism. New Essays in Metaphysics*, edited by T. Goldschmidt and K. L. Pearce, Oxford: Oxford University Press, 200–216.

Pepper, J. W. and Herron, M. D. (2008), "Does Biology Need an Organism Concept?", in *Biological Review*: 83, 621–627.

Queller, D. and Strassmann, J. (2009), "Beyond Society: the Evolution of Organismality", in *Philosophical Transactions of the Royal Society B*: 364, 3143–3155.

Simons, P. (2000), "Continuants and Occurrents", in *The Aristotelian Society Supplementary*: 74, 78–101.

Simons, P. (2005), "The Ties that Bind: What Holds Individuals Together", in *Substanz. Neue Überlegungen zu einer klassischen Kategorie des Seienden*, edited by K. Trettin, Frankfurt/M: Klostermann, 229–243.

Simons, P. (2018), "Processes and Precipitates", in *Everything Flows*: *Towards a Processual Philosophy of Biology*, edited by D. J. Nicholson and J. Dupré, Oxford: Oxford University Press, 49–60.

Van Inwagen, P. (1990), *Material Beings*, abbreviated by *MB*, Ithaca: Cornell University Press.

Wilson, J. A. (2000), "Ontological Butchery: Organism Concepts and Biological Generalisations", in *Philosophy of Science*: 67, 301–311.

William Jaworski

Peter van Inwagen and the Hylomorphic Renaissance

Abstract: Peter van Inwagen is well known for his groundbreaking work on composition. Less well known is the way that work has contributed to the rehabilitation of hylomorphism, the view that some individuals – paradigmatically living things – are composed of physical materials with a specific organization or structure. I explain how van Inwagen's ideas provide the basis for a hylomorphic theory that dovetails with current work in biology and biological subdisciplines such as neuroscience. That theory differs from van Inwagen's in an important respect: it is committed to the existence of functional parts such as eyes, hearts, and brains. I argue that this commitment is a straightforward implication of van Inwagen's claim that the details about what lives are and what characteristics they have need to be supplied empirically, for actual empirical work on living things posits functional parts. I defend this implication against its detractors, and argue that a commitment to functional parts enables hylomorphists to respond more effectively to the objections that have been levied against van Inwagen's own view.

1 Van Inwagen on composition

Peter van Inwagen is well known for his groundbreaking work on composition. Lesser known is the way his work has contributed to the rehabilitation of hylomorphism. Hylomorphism's basic idea, stated very roughly, is that some things are composed of matter with a specific organization or structure. A human being is not composed of physical materials structured in any way whatsoever, but physical materials structured in a very specific way. In some cases, a thing's structure (its form) is something static, like the relatively unchanging spatial arrangements of atoms in a crystal, but in other cases – by far the more interesting ones – the structure comprises dynamic interactions among an individual's components. The configurations of matter and energy that make human beings and other complex living things what they are cannot be characterized apart from the dynamic interactions among their various organ systems, along with their component organs, tissues, cells, and the molecules, atoms, and fundamental physical materials ultimately composing them – in short, they cannot be characterized apart from the kinds of things van Inwagen calls 'lives'.

In what follows, I explain how van Inwagen's ideas provide the basis for a hylomorphic theory that dovetails with current work in biology and biological subdisciplines such as neuroscience. That theory differs from van Inwagen's in a few

https://doi.org/10.1515/9783110664812-014

respects. One of them is its commitment to functional parts such as eyes, hearts, and brains. Van Inwagen denies that there are such parts. I nevertheless argue that their existence is commensurate with his claim that the details about what lives are and what characteristics they have need to be supplied empirically since actual empirical work on living things posits functional parts. I defend this implication against some objections, and argue that it allows for an effective response to objections that have been levied against his view.

Van Inwagen presents his view as an answer to the Special Composition Question: under what conditions do many things compose one thing? He answers that composition happens exactly if the activities of physical particles constitute a life. This implies that something qualifies as a part only if it is caught up in a life – an expression he borrows from the biologist J. Z. Young (1971).[1] Van Inwagen's descriptions of lives stay largely at the level of metaphor and analogy. The reason is that providing the literal details about what lives are and what characteristics they have is, he thinks, a job for biologists (Van Inwagen (1990), p. 84). He does nevertheless offer some general characteristics. Lives, he says, are self-maintaining events like flames and waves except that unlike flames and waves they are *well-individuated* and *jealous*. Flames are not as well-individuated as lives, van Inwagen argues, and unlike waves, lives are jealous: two waves can be composed of the activities of the same water molecules, but two lives cannot.

Composition does not happen apart from lives; composite beings are all living things. Consequently, on van Inwagen's view, there are only two kinds of material beings: mereological simples (material beings with no proper parts) and living things. Van Inwagen's view is well known for implying the Denial, the claim that many objects in a commonsense ontology do not exist, including artifacts and natural bodies such as mountains and planets. There is no table occupying the

1 Van Inwagen explains with an example: "Alice drinks a cup of tea in which a lump of sugar has been dissolved. A certain carbon atom … is carried along with the rest of the sugar by Alice's digestive system to the intestine. It passes through the intestinal wall and into the bloodstream, whence it is carried to the biceps muscle of Alice's left arm. There it is oxidized in several indirect stages (yielding in the process energy … for muscular contraction) and is finally carried by Alice's circulatory system to her lungs and there breathed out as a part of a carbon dioxide molecule … Here we have a case in which a thing, the carbon atom, was … caught up in the life of an organism, Alice. It is … a case in which a thing became however briefly, a part of a larger thing when it was a part of nothing before or after …" (Van Inwagen (1990), pp. 94–95).

region of space before me – no single, unified individual. There are instead many physical particles spatially arranged table-wise.

Why limit the Denial to artifacts and natural bodies? Why not deny that all composites exist? Van Inwagen's most compelling answer is that living things have non-redundant causal powers that other alleged composites lack. The activities attributed to artifacts and natural bodies can be understood as disguised cooperative activities performed by simples. The chair, the mountain, and the planet don't do anything that cannot be exhaustively described and explained by appeal to the activities of mereological simples, but according to van Inwagen, not all activities are like this. Living things can do things that cannot be done by simples alone but only by composite individuals.

Those individuals include only single cells and multicellular organisms on van Inwagen's account. There are no functional parts such as eyes, hearts, and brains. The reason, Van Inwagen (1981) argues, is that there is no principled basis for dividing an organism into functional parts as opposed to, say, thirds, or halves, or parts of some other sort. Functional divisions do not correspond to joints in the natural world, but only to our peculiar descriptive interests. We can easily imagine a race of beings with descriptive interests different from ours who would divide humans into parts in a different way. Functional parts are thus arbitrary. Arbitrary parts, however, do not exist, for their existence would generate philosophical problems such as the body-minus problem (Wiggins (1968); Burke (1994)). Suppose Descartes loses his left leg at time *t*. The following all seem to be true:

> (1) Descartes exists before *t*, and also (2) a proper part of him, his left leg, exists before *t*. (3) If Descartes' left leg exists before *t*, then another part of him, Descartes-minus-his-left-leg (call it 'D-minus') exists before *t*. Imagine, however, that (4) Descartes survives the amputation of his left leg at *t*. Since, (5) D-minus also survives the amputation of Descartes' left leg at *t*, it seems that after *t* Descartes must be identical to D-minus. The reason is that (6) after *t* Descartes and D-minus have the same size, shape, position, orientation, attitude, mass, velocity, and color, but (7) two objects cannot have all these characteristics in common; hence, Descartes and D-minus must be identical.[2] The problem is that identifying Descartes and D-minus appears to violate the axioms of identity. Because Descartes used to have two legs but D-minus didn't, for instance, it appears to violate Leibniz's law. Moreover, as van Inwagen argues, the identity of Descartes and D-minus also appears to violate the transitivity of identity.

2 Burke (1994) and Wiggins (1968) appeal to different principles here: Burke, to the principle that two things cannot share all the same parts at the same time, and Wiggins to the principle that two things of the same sort cannot occupy the same place at the same time.

Van Inwagen's solution is to reject claim (2): if there is no such thing as Descartes' left leg, then it no longer follows that there is such a thing as D-minus, and if there is no such thing as D-minus, the body-minus problem never gets off the ground. What is true of Descartes' left leg is true mutatis mutandis of other functional parts, and indeed, of proper parts in general other than mereological simples and single cells. Why the exception for single cells? Their lives are subordinate to the lives of the organisms to which they belong, van Inwagen tells us – an idea to which we'll return later.

2 Hylomorphism

With this outline of van Inwagen's view in place, let's turn to hylomorphism. The past fifteen years have witnessed a renaissance of hylomorphic thinking. Several metaphysicians have articulated hylomorphic theories including Fine (1999, 2008), Johnston (2006), Oderberg (2007), Koslicki (2008), Rea (2011), Marmodoro (2013), Koons (2014), Evnine (2016), and Jaworski (2014, 2016). Hylomorphism claims that form or structure is a basic ontological and explanatory principle. Some individuals, living things, consist of materials that are structured or organized in various ways. You and I are not mere quantities of physical materials; we are individuals composed of physical materials with a certain organization or structure. That structure is responsible for us being and persisting as humans, and it is responsible for us having the developmental, metabolic, reproductive, perceptive, and cognitive capacities we have.

When people hear the term 'structure' they often think of something static such as the relatively unchanging spatial relations among atoms in a crystal.[3] But hylomorphists don't view structure so narrowly. Although we're free to call the sum of spatial relations among something's parts a 'structure' in some sense of the term, hylomorphic structures – the kind that, say, distinguish living things from nonliving ones – are not static spatial relations, but dynamic patterns of environmental interaction. They comprise programmatic sequences of changes over time, and often involve different kinds of changes under different kinds of conditions. It is because of their dynamic structures – their abilities to impose their structures on incoming matter and energy – that composite individuals (paradigmatically

3 In fact, even some contemporary hylomorphists use the term 'structure' this way. Oderberg (2014), p. 177, is an example.

living things) persist one and the same through the constant influx and efflux of matter and energy that characterize their interactions with the wider world.

It's also worth noting that the hylomorphic notion of structure is not the same as others that have appeared in the literature. It is not the same, for instance, as the notion of structure that has been operative in discussions of grounding in metaphysics (Schaffer (2009); Sider (2012)), or the notion operative in debates about scientific realism (Worrall (1989); Ladyman and Ross (2007)), or the notion David Chalmers sometimes employs when he speaks of structure and dynamics (Chalmers (2002), p. 258).

What, then, is hylomorphic structure? Because the notion of structure is a primitive within a hylomorphic framework, it cannot be defined in terms of anything more basic; rather, like the notions of necessity and possibility in modal actualist theories (Plantinga (1979)), it can only be defined by appeal to the roles it plays within the framework, and then illustrated with examples.[4] I've summarized those roles with some slogans:

> *Structure matters*: it operates as an irreducible ontological principle, one that accounts at least in part for what things essentially are.
>
> *Structure makes a difference*: it operates as an irreducible explanatory principle, one that accounts at least in part for what things can do, the powers they have.
>
> *Structure counts*: it explains the unity of composite things, including the persistence of one and the same living individual through the dynamic influx and efflux of matter and energy that characterize many of its interactions with the wider world.

A simple example helps illustrate the hylomorphic notion of structure; we can call it *the squashing example*. Suppose we put Dean in a strong bag – a very strong bag since we want to ensure that nothing leaks out when we squash him with several tons of force. Before the squashing the contents of the bag include one human being; after they include none. In addition, before the squashing the contents of the bag can think, feel, and perceive, but after the squashing they can't. What explains these differences in the contents of the bag pre-squashing and post-squashing? The physical materials (whether particles or stuffs) remain the same – none of them leaked out. Intuitively we want to say that what changed was the way those materials were structured or organized. That organization or structure was responsible for there being a human before the squashing, and for that human

4 As Kit Fine says, the only way of defining a framework's basic concepts is, "to specify the principles by which [they are] governed" (Fine (2008), p. 112).

having the capacities it had. Once that structure was destroyed, there no longer was a human with those capacities. Structure is thus a basic ontological principle: it concerns what things there are. It is also a basic explanatory principle: it concerns what things can do.

Grene (1972), pp. 409-410, once noted that the hylomorphic notion of structure is close to the notion of organization that many biologists appeal to. Here is one example taken from a popular college-level biology textbook – note the references to organization, order, arrangement, and related concepts:

> Life is highly organized into a hierarchy of structural levels … Biological order exists at all levels. … [A]toms … are ordered into complex biological molecules … the molecules of life are arranged into minute structures called organelles, which are in turn the components of cells. Cells are [in turn] subunits of organisms… The organism we recognize as an animal or plant is not a random collection of individual cells, but a multicellular cooperative… Identifying biological organization at its many levels is fundamental to the study of life. … With each step upward in the hierarchy of biological order, novel properties emerge that were not present at the simpler levels of organization . … A molecule such as a protein has attributes not exhibited by any of its component atoms, and a cell is certainly much more than a bag of molecules. If the intricate organization of the human brain is disrupted by a head injury, that organ will cease to function properly. … And an organism is a living whole greater than the sum of its parts. … [W]e cannot fully explain a higher level of order by breaking it down into its parts. (Campbell (1996), pp. 2–4)

This passage suggests that the way things are structured or organized plays an important role in them being the kinds of things they are, and in explaining the kinds of things they can do. Organization matters and organization makes a difference: it operates as an ontological and explanatory principle.

The same appears to be true of van Inwagen's lives: *Lives matter* on van Inwagen's view; they are ontological principles: whether the *x*s constitute a life makes a difference to whether a composite individual exists. Likewise, *lives make a difference*; they are explanatory principles: living beings can do things that cannot be exhaustively described and explained using the conceptual resources used to describe and explain the materials composing them (Van Inwagen (1990), pp. 122, 180). Finally, *lives count*; they operate as principles of unity (Van Inwagen (1990), p. 121) and persistence (Van Inwagen (1990), pp. 145, 148): what binds the simples that compose me into a single being is that their activity constitutes a life, and what enables me to persist through changes in those simples is the persistence of that life. Because van Inwagen's lives play these roles it is easy to use his view of composition as a basis for understanding the hylomorphic view.

Configuring materials and being composed of materials are co-foundational concepts on the hylomorphic view, just as having a life and being composed of simples are co-foundational concepts on van Inwagen's. Likewise, just as van In-

wagen restricts composition to living things, hylomorphists restrict it to structured things in general. According to hylomorphists, composition occurs when and only when an individual configures materials.

Structured individuals are *emergent* individuals on the hylomorphic view: there are empirically-describable conditions that are sufficient to bring into existence a new structured individual where previously no such individual existed. Suppose, for instance, that b_1, b_2,..., b_n are physical particles or materials of some sort. On the hylomorphic view there are changes the *b*s can undergo which will result in there being a new individual, a, which is composed of the *b*s. In the natural course of human events, for instance, changes of this sort regularly happen in utero: physical materials that didn't compose a human organism at time t_1 come by a series of changes to compose a human organism at time t_2. A new human individual comes to exist where previously no such individual did.

Once a structured individual comes into existence it is continuously engaged in configuring materials, and the materials it configures are precisely those that compose it. The individual a comes into existence exactly with the start of its configuring activity – exactly when that configuring activity begins. When it comes to characterizing that activity, hylomorphists can adopt most of what van Inwagen says about lives, at least when it comes to the configuring activities of living things, the paradigmatic structured individuals. My life is identical to my configuring various fundamental physical materials at various times – an event that has the characteristics van Inwagen attributes to lives, and that has many other characteristics it is the business of the biological sciences to describe.

An individual living thing does not configure the same materials for very long; the materials composing it are in constant flux. If a's existence commences with its configuring the *b*s, it will not take long for it to exchange some of the *b*s for other things. Yet despite this, a maintains itself one and the same through these changes on account of its ongoing configuring activity. That activity is what unifies various materials into a single individual, both synchronically and diachronically, just as lives do on van Inwagen's account.

How do we know in any given case whether a quantity of physical materials compose a unified whole? How do we know, for instance, that the physical materials located in this region of space actually compose a human being, that they are not instead diverse materials that are merely spatially arranged human-wise and that do not compose a unified whole at all? Here too hylomorphists can take a cue from van Inwagen: structured individuals have non-redundant causal powers that mere spatial arrangements of physical materials do not.[5] Our initial impres-

5 This response implies a certain meta-mereology which Kathrin Koslicki describes as follows: "I

sion that *a* is a structured whole composed of the *b*s could be accurate or inaccurate. Determining which is a matter of determining whether *a* has powers not had by the *b*s, and determining whether this is the case is a matter of determining whether descriptions and explanations of *a*'s behavior are reducible to descriptions and explanations of the *b*s' behavior.

The hylomorphic view clearly implies a kind of property pluralism since structured individuals have properties of at least two sorts: properties due to their structures (or their integration into individuals with structures), and properties due to the materials composing them independent of the way those materials are structured. This is illustrated by the squashing example considered earlier. Dean's powers to think, feel, and perceive are clearly structure-dependent properties: destroying his structure destroys those properties. By contrast, the squashed contents of the bag have the same mass that Dean has despite losing Dean's human structure. Mass is a structure-independent property.

Some philosophers and biologists call the new properties of structured systems *emergent properties*. Emergent properties have three characteristics:

> They are first-order properties, not higher-order ones; that is, they are not logical constructions with definitions that quantify over other properties; they are rather powers in their own right.
>
> They are not epiphenomenal, but make distinctive causal or explanatory contributions to the behavior of the individuals having them.
>
> They are possessed by an individual on account of its organization or structure.

Notice: it is not a characteristic of emergent properties (at least not on the hylomorphic view) that they are generated or produced by lower-level systems. As a result, hylomorphists do not need an account of how lower-level systems generate emergent properties. Emergent properties are due to something's structure, and structure is a basic principle on the hylomorphic view; it is not generated by anything else.

---

take the mereologist's job to be to devise an appropriate conception of parthood and composition which accurately reflects the conditions of existence, spatio-temporal location and part/whole structure of those objects to which we take ourselves to be already committed as part of the presupposed scientifically informed, commonsense ontology. The question of which kinds [of objects] there are I take to be… answered [not] by the mereologist proper, but by the ontologist at large, in conjunction with… science and common sense, which… have something to contribute to the question, 'What is there?'" (Koslicki (2008), p. 171).

Hylomorphic structures, then, are powers to configure (or organize, order, or coordinate) things. What sets them apart from other powers is that they cannot exist unmanifested. They are manifested essentially. Structured individuals are essentially and continuously engaged in configuring the materials that compose them. I configure the materials that compose me, and you configure the materials that compose you. Our continuous structuring activities explain our respective unities and our persistence through the dynamic influx and efflux of matter and energy that characterize our interactions with the surrounding world. This is what it means to say that *structure counts*: it explains the unity of composite things, and it does so in a way that resembles van Inwagen's account of composition. For van Inwagen, *x* is a part of *y* exactly if *x* is caught up in the life of *y*. For the hylomorphist, *x* is a part of *y* exactly if *y* is a structured individual, and *x* is among the things that *y* configures.

3 Atomism, universals, non-living wholes, and functional parts

There are nevertheless four noteworthy differences between the hylomorphic view and van Inwagen's. First, van Inwagen is a committed atomist; he claims that fundamental physical materials are particulate. Hylomorphists are not committed to this. They assume that fundamental physical entities are capable of composing structured wholes, but they do not take a further stand on the natures of those entities. They instead leave it to the relevant empirical disciplines to tell us what their natures are.[6] I've argued in detail elsewhere (Jaworski (2016)) that this difference helps address gunk-based objections to van Inwagen's view such as those advanced by Sider (1993) and Zimmerman (2003).

Second, van Inwagen is committed to the existence of universals. I've argued elsewhere, however, that hylomorphic structures are best conceived as tropes (Jaworski (2016)). My structure and your structure – the ongoing configuring activities in which we engage – are best conceived as particulars that are numerically distinct from each other. They count as structures of the same type – as human structures, say – only insofar as they closely resemble each other.

6 In order to express this neutrality about the nature of fundamental physical entities, I've used to the term 'materials' to refer to them since the term 'materials' can be applied both to discrete individuals and to continuous stuffs ('building materials', for instance, can be applied both to individual timbers and nails, and to stuffs such as glue, metal, and wood).

Third, van Inwagen limits composition to living things, but hylomorphism is open to the possibility that there might be structured individuals of other sorts. Van Inwagen claims, for instance, that atoms and molecules do not exist; they are virtual objects which are virtually composed of mereological simples. Hylomorphists are free to take the same stance as van Inwagen, but two more options are available to them. They can claim that atoms and molecules are structured individuals in their own right distinct from living things and their parts, or they can claim that atoms and molecules are parts of living things and the atom- and molecule-wise arrangements of fundamental physical materials we find outside organisms are virtual objects in van Inwagen's sense. Which stance they ultimately take depends on broadly empirical considerations. If atoms and molecules are discovered to have powers distinct from those that can be exhaustively described and explained by appeal to fundamental physical materials alone, then there are grounds for claiming that they are not mere aggregates of fundamental physical materials, but distinctive individuals in their own right.

Finally, the hylomorphic view of parts is less revisionary than van Inwagen's. Van Inwagen denies that there are functional parts such as eyes, hearts, and kidneys. Hylomorphists, by contrast, accept that there are such parts. Their reasons are again broadly empirical: our best descriptions and explanations of human behavior postulate parts like these, and this gives us good reason to think these parts exist. It is worth considering the argument for this claim in greater detail.

There are multiple ways of dividing things into parts. A hammer can be exhaustively decomposed into functional parts such as the head, the claw, and the handle – parts that contribute to the hammer's overall operation. It can also be exhaustively decomposed into spatial parts such as the top third, the middle third, and the bottom third, or into the spatial parts of some other sort. Different principles of part identity and individuation yield different inventories of parts, all of which may be consistent with principles of formal mereology such as the transitivity and asymmetry of proper parthood. Because there are many different principles for part identity and individuation, whenever we want to consider the parts that an individual has we need to determine which principle is (or which principles are) best suited to our descriptive and explanatory ends. If we accept a broad naturalism in matters of ontology, then empirical adequacy is an important criterion for making that determination.

Ontological naturalism implies that when it comes to determining what exists, empirical investigation – paradigmatically science – is our best guide. We can think of this as the conjunction of a Quinean thesis about ontological commitment with a broad empiricism: we are committed to all the entities postulated by our best descriptions and explanations of reality, and those descriptions and explanations derive from empirical sources such as the natural and social sci-

ences. Consequently, when it comes to determining what parts exist, and what principles of part identity and individuation carve the world at its joints, ontological naturalism implies that we should endorse principles that reflect our best empirical descriptions and explanations of their behavior. Roughly, if principle *P* does a better job enabling us to describe and explain the behavior of *K*s than principle *P**, then we have good reason to accept that *K*s have parts that are identified and individuated by *P* instead of *P**.

Van Inwagen tells us that describing what lives are is the business of biology, just as describing what metals are is the business of chemistry (Van Inwagen (1990), p. 84). When we turn to biology, however, we find descriptions of living things that posit functional parts. Although humans can be divided into parts in many different ways, biologists, neuroscientists, psychologists, and others are primarily interested in dividing them along functional lines (Bechtel (2007, 2008); Craver (2007)).[7] Perhaps the best example of this is the method *functional analysis* (also called 'mechanical decomposition' or 'functional decomposition').[8] It analyzes the activities of complex systems into simpler subactivities performed by simpler subsystems (Fodor (1968); Cummins (1975); Lycan (1987), Chapter 4; Bechtel and Richardson (1993); Glennan (2002, 2017); Bechtel (2007, 2008); Craver (2007)).[9] Functional analysis of a complex human activity – running, say – reveals a circulatory subsystem responsible for supplying oxygenated blood to the muscles. Analysis of that subsystem reveals a component responsible for pumping the blood – a heart. Analysis of the heart's pumping activity shows that it is composed of muscle tissues that undergo frequent contraction and relaxation. These activities can be analyzed into the subactivities of various cells, and these in turn into the operations of various organelles. This analytic process continues until something is discovered to have a property or to engage in an activity not on

7 Carl Craver (2007), Chapter 5, calls purely spatial parts 'pieces' and parts in the functional sense 'components'. John Heil (2003), p. 100, also suggests something like the distinction between merely spatial parts and parts of other sorts, which he calls 'substantial parts'.

8 The term 'functional analysis' is due to Cummins (1975). Bechtel (2008) calls it 'mechanistic decomposition' or 'functional decomposition'. Craver (2007) subsumes it under the heading of 'mechanistic explanation'. He takes Cummins's notion of functional analysis to be the exemplar of what he calls the 'systems tradition', but argues that Cummins fails to provide an adequate account of mechanistic explanation. He thus distances himself from the term 'functional analysis'.

9 Elsewhere I've discussed functional analysis in greater detail. I've argued among other things that it does not correspond to the notion of function that is operative in discussions of functionalism in the philosophy of mind (including teleological functionalism), and that it does not imply a commitment to reductionism in the sense of one theory taking over the descriptive and explanatory jobs of another (Jaworski (2011, 2012)).

account of the activities of some lower-level subsystems but as an unanalyzable matter of fact.

Functional analysis is important because it provides a basis for understanding the kinds of parts postulated by descriptions and explanations in the biological sciences – the sciences charged with the task of telling us what lives are. They tell us that the parts of living things are subsystems or components – things that contribute in empirically-specifiable ways to the activities of the wholes to which they belong. So if we accept a broad ontological naturalism of the sort I've described, then we ought to accept that there are functional parts.

This should be reckoned a good thing. Many of van Inwagen's critics balk at the Denial. A view that can accommodate functional parts softens the Denial's implications. Hylomorphism enables us to countenance the kinds of parts postulated by work in the biological sciences, and to the extent that those parts overlap with parts in a commonsense ontology, it helps us accommodate the deliverances of common sense as well.

Of course, if functional parts generate insoluble philosophical problems we'll have to reexamine our commitment to them. But do they? More to the point: are the alleged problems with functional parts best solved in the way van Inwagen suggests? Consider again the body-minus problem. His solution is to deny that Descartes has any proper parts other than fundamental physical particles and individual cells, which implies that claim (2) is false. If there is no such thing as Descartes' left leg, then it no longer follows that there is such a thing as D-minus, and if there is no such thing as D-minus, the body-minus problem never gets off the ground. Since hylomorphists endorse functional parts, this solution is not open to them. They look instead to reject claim (3); they deny that the existence of functional parts implies the existence of functional part-complements. If left legs exist, it does not follow that left-leg-complements such as D-minus do.

Given some plausible assumptions, moreover, they also reject claim (5): there are good reasons to think functional part-complements do not exist, and if that is true, then it is false to claim that D-minus exists after the amputation – D-minus doesn't exist at all.

Van Inwagen defends (3) on the grounds that postulating legs but not leg-complements would be arbitrary. After all, Descartes' left leg and D-minus comprise exactly the same kinds of physiological processes – cellular respiration, cellular replication, protein synthesis, and so forth. What principled basis could there be for claiming that the physiological processes in Descartes' left leg constitute a life while those in D-minus do not? The burden for any solution to the body-minus problem that rejects (3) is to provide a principled basis for distinguishing legs, hands, and hearts, from leg-, hand-, and heart-complements.

One possible basis is the ontological naturalism mentioned a moment ago. It suggests that not all principles of part identity and individuation are created equal. Some ways of dividing organisms into parts are privileged, in particular the ways implied by our best empirical descriptions and explanations. If those descriptions and explanations divide organisms into functional parts but not functional part-complements, then we have good prima facie reason to think the former parts exist but not the latter.

Do our best empirical descriptions and explanations divide organisms into functional parts in fact? The answer appears to be yes: biologists, neuroscientists, psychologists and others who are concerned with giving empirical descriptions and explanations of living behavior tend to view organisms as complex systems with parts that are individuated functionally in ways discovered through functional analysis. This gives us good prima facie reason to think that our best empirical descriptions and explanations postulate functional parts like eyes, legs, and hearts.

They also give us good reason to think that those descriptions and explanations do not postulate eye-, leg-, and heart-complements like D-minus. Actual biological descriptions and explanations individuate parts in ways that are too fine-grained to accommodate parts like D-minus – parts that would include within their boundaries segments of many diverse subsystems: the circulatory system, the reproductive system, the visual system, and so on. It is possible that biologists, neuroscientists, and others might eventually discover that human behavior is described and explained more effectively using radically different principles of part identity and individuation, but in the absence of compelling reasons to think this will happen, our best biological science provides some reason to think that there are no such parts as D-minus.

If there are hands, legs, and hearts, but not hand-, leg-, and heart-complements, then claim (3) of the body-minus problem is false. The existence of functional parts like Descartes' left leg does not imply that there are functional part-complements like D-minus. The foregoing considerations also give us good reason to think that functional part-complements like D-minus do not exist. As a result, there can be no question of whether Descartes is identical to D-minus, for there is no D-minus. This implies that claim (5) of the body-minus problem is false as well. Claim (5) says that D-minus survives the amputation of Descartes' left leg. If we assume – plausibly – that a statement of the form '*x* survives *y* at *t*' implies that *x* exists after *t*, then claim (5) is false if D-minus does not exist.

Hylomorphists thus have a way of dealing with one of van Inwagen's arguments against the existence of functional parts. But van Inwagen suggests a second argument as well. Lives are jealous, he says – so jealous that it is impossible for two lives to overlap unless one of them is subordinate to the other.

The only case in which this can occur, he says, is a case in which the life of an individual cell is subordinate to the life of a multicellular organism (Van Inwagen (1990), p. 89). Since eyes, hearts, and similar functional parts are not individual cells, they cannot be subordinate to the life of a multicellular organism. But neither are they multicellular organisms themselves; if they exist at all, they must be proper parts of multicellular organisms. Consequently, eyes, hearts, and similar functional parts must not exist.

Hylomorphists can respond by challenging the premise that the only cases in which lives are subordinate are cases involving individual cells. Van Inwagen's assumption seems to be that if there were subordinate lives other than those of individual cells they would undermine the jealousy of lives. Yet it is not entirely clear why we should accept this assumption, nor why the jealousy of lives is not supposed to be undermined by the subordinate lives of individual cells and yet is supposed undermined by the subordinate lives of other things.

One plausible reason might concern the unifying roles that lives are supposed to play. Recall that *lives count* on van Inwagen's view; a life is supposed to explain why diverse things compose a single unified whole. The larger the number and variety of things involved, the more unifying work the notion of a life must be pressed into service to perform. Someone might feel that lives could plausibly be asked to unify diverse cells and physical particles, but that it would take things too far to ask them to unify multicellular tissues, organs, and organ systems in addition.

Hylomorphists could respond by elaborating the notion of subordination to include a notion of subordinate unity: a multicellular organ enjoys a unity analogous to that of a multicellular organism. It is a subpocket of order within a larger ordered whole. To use van Inwagen's analogy between an organism and a storm, an organism would be like a large hurricane that comprises smaller tornadic vortices within itself. Each smaller pocket of order – or suborder – is unified in a way that enables it to perform the subactivity that qualifies it as a proper part of a whole.

This subordinate unity might explain why so-called "severed parts" sometimes display behaviors similar to the behaviors of parts *in situ*. A heart removed from the chest cavity continues to beat (or more precisely, the materials that used to compose a heart continue to change position relative to each other in a manner that closely resembles the beating of a heart). This lifelike behavior can be explained as a vestige of the unity those materials once possessed when they composed a heart. That vestige of unity can be used to explain in turn why it remains possible for those materials to recompose a heart when they are transplanted to a new individual even if the same materials are not transplantable in the same way if smashed or incinerated. The latter processes disrupt the vestige of unity those

materials enjoyed as parts, but if that vestige remains intact, then it remains possible for the materials to take up once again the kind of role they formerly played.

Suppose, then, that parts enjoy this kind of subordinate unity. Because the unity is subordinate to the unity of the wholes to which they belong, their existence should not compromise the jealousy of lives any more than the existence of individual cells does. Nothing stands in the way, therefore, of hylomorphists countenancing the existence of functional parts.

4 The Brainstem objection

Critics might nevertheless insist that hylomorphists have failed to solve the body-minus problem. Suppose that instead of D-minus, which admittedly stands little empirical chance of qualifying as a functional part, we instead take something that uncontroversially qualifies: Descartes' brainstem. Why the brainstem? Eric Olson (1996) has argued that there is good empirical reason to suppose that the brainstem is the functional part with which at the barest minimum a human can survive: a human survives if and only if his or her brainstem does. Call this the 'brainstem survival thesis'. We can now reformulate the original body-minus problem using Descartes' brainstem, B, in place of D-minus:

> (1) Descartes exists before *t*, and (2) a proper part of him, namely his brainstem, B, exists before *t*. Due to some catastrophe, however, every proper part of Descartes except B and its proper parts is destroyed at time *t*. (3) If B is the bare minimum functional part with which a human animal can survive, then Descartes survives the catastrophe. So Descartes exists after *t*. But (4) B also survives the catastrophe, so it too exists after *t*. Now, it seems that after *t* Descartes must be identical to B since (5) after *t* Descartes and B have the same size, shape, position, orientation, attitude, mass, velocity, and color, but (6) two objects cannot have all these characteristics in common. Hence, Descartes and B must be one, yet identifying Descartes and B would violate Leibniz's law since Descartes used to have arms and legs, but B did not. There are thus good reasons to think both that Descartes and B are identical, and that they are distinct.

One advantage of van Inwagen's solution is that it elegantly solves both the earlier body-minus problem and this reformulated problem the same way. In both cases, it denies that Descartes has a proper part, either his left leg or his brainstem. Hylomorphists, however, are not free to deny the existence of parts postulated by our best empirical descriptions and explanations of human behavior, and based on what's been said, that includes functional parts such as brainstems. Hylomorphists must therefore solve the reformulated problem a different way.

They target claim (4). According to them 'B' designates a proper part of Descartes that immediately prior to t was composed of objects f_1, f_2,..., f_n (the cells, organelles, molecules, and so forth that would be revealed through functional analyses of B's activities before *t*). After time *t*, however, the *f*s no longer compose a proper part of Descartes; they compose only Descartes himself. What had been a proper part of Descartes, his brainstem, no longer exists even though Descartes does.

This solution implies that the brainstem survival thesis is false, but hylomorphists can accommodate something like it. Suppose that instead of the brainstem, it is rather the *f*s that play the role of being the minimum functional core that Descartes needs to survive. Descartes will continue to exist as long as the *f*s continue to contribute to his overall operation as they have done hitherto, but if any of them is damaged or destroyed, and is incapable of performing its contributing subactivity at a time, then Descartes will cease to exist at that time. Collectively, then, the *f*s constitute a core of functional components that Descartes needs to exist, and that are sufficient to enable him to exist. They are in this sense the barest minimum functional parts with which Descartes can survive. The *f*s composed Descartes' brainstem prior to *t*, and as a result, someone might be tempted to say that Descartes' brainstem is the bare minimum Descartes needs to survive, but strictly speaking this is false. What is true instead is that the *f*s are the bare minimum that Descartes needs to survive, and the *f*s needn't compose a brainstem. Descartes thus survives the catastrophe along with the proper parts which used to compose his brainstem but which now compose only him.

Critics might complain that it is implausible to suppose that Descartes' brainstem would disappear even though all the parts that previously composed it remain where they are. There are at least two things hylomorphists can say in response. First, the objection appears to be tacitly committed to the idea that the intrinsic properties and the spatial arrangement of the *f*s are sufficient for those things to compose a brainstem. But a functional principle of part identity and individuation suggests otherwise. For the *f*s to compose something, they must contribute to the activities of the whole they compose. Suppose, then, that brainstems are essentially things that coordinate the activities of diverse organ systems (van Inwagen himself appears to endorse a principle like this (Van Inwagen (1990), pp. 177–178)). When Descartes is whittled down to the *f*s, he no longer has any diverse organ systems to coordinate. In that case, however, it is not implausible to claim that he does not have a brainstem among his proper parts.

Second, hylomorphists can argue that even if it is implausible to suppose that Descartes' brainstem ceases to exist when he is whittled down, it is nevertheless no more implausible than the alternatives. It is no more implausible, for instance, than claiming that Descartes' brainstem never existed at all. But if this solution

to the body-minus problem is no more implausible than the alternatives, then the charge of implausibility loses its force.

I've argued in favor of the hylomorphic view in greater detail elsewhere (Jaworski (2016)). There's much more to be said about it – including how it handles other well-known problems of composition. I've had to compress many points and avoid discussing others altogether, but I do hope I've said enough at least to introduce the view, and to clarify to some extent how van Inwagen's groundbreaking work has contributed to its rehabilitation.

Bibliography

Bechtel, W. (2007), "Reducing Psychology while Maintaining its Autonomy via Mechanistic Explanations", in *The Matter of the Mind*, edited by M. Schouten and H. L. de Jong, Blackwell Publishing, 172–198.

Bechtel, W. (2008), *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience*, New York: Routledge.

Bechtel, W., and Richardson, R. (1993), *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*, Princeton, NJ: Princeton University Press.

Burke, M. (1994), "Dion and Theon: An Essential Solution to an Ancient Puzzle", in *Journal of Philosophy*: 91, 129–139.

Campbell, N. A. (1996), *Mental Mechanisms: Biology*, 4th Edition, The Benjamin/Cummings Publishing Company, Inc..

Chalmers, D. J. (2002), "Consciousness and Its Place in Nature", in *Philosophy of Mind*, edited by D. Chalmers, Oxford: Oxford University Press, 247–272.

Craver, C. F. (2007), *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*, New York: Oxford University Press.

Cummins, R. (1975), "Functional Analysis", in *Journal of Philosophy*: 72, 741–764.

Dennett, D. C. (1978), *Brainstormse*, Cambridge, MA: Bradford Books.

Evnine, S. J. (2016), *Making Objects and Events: A Hylomorphic Theory of Artifacts, Actions, and Organisms*, Oxford: Oxford University Press.

Fine, K. (1999), "Things and Their Parts", in *Midwest Studies in Philosophy*: 23, 61–74.

Fine, K. (2008), "Form and Coincidence", in *Proceedings of the Aristotelian Society*, Supplementary: 82, 101–118.

Fodor, J. (1968), "The Appeal to Tacit Knowledge in Psychological Explanation", in *Journal of Philosophy*: 65, 627–640.

Glennan, S. (2002), "Rethinking Mechanistic Explanation", in *Philosophy of Science*: 69, 342–353.

Glennan, S. (2017), *The New Mechanical Philosophy*, Oxford: Oxford University Press.

Grene, M. (1972), "Aristotle and Modern Biology", in *Journal of the History of Ideas*: 33, 395–424.

Heil, J. (2003), *From an Ontological Point of View*, Oxford: Clarendon.

Jaworski, W. (2011), *Philosophy of Mind: A Comprehensive Introduction*, Malden, MA: Wiley-Blackwell.

Jaworski, W. (2012), "Powers, Structures, and Minds", in *Powers and Capacities in Philosophy*: *The New Aristotelianism*, edited by J. Greco and R. Groff, Routledge, 145–171.
Jaworski, W. (2014), "Hylomorphism and the Metaphysics of Structure", in *Res Philosophica*: 91, 179–201.
Jaworski, W. (2016), *Structure and the Metaphysics of Mind*: *How Hylomorphism Solves the Mind-Body Problem*, Oxford: Oxford University Press.
Johnston, M. (2006), "Hylomorphism", in *Journal of Philosophy*: 103, 652–698.
Koons, R. (2014), "Staunch vs. Faint-hearted Hylomorphism: Toward an Aristotelian Account of Composition", in *Res Philosophica*: 91, 151–178.
Koslicki, K. (2008), *The Structure of Objects*, Oxford: Oxford University Press.
Ladyman, J., and Ross, D. (2007), *Every Thing Must Go*: *Metaphysics Naturalized*, Oxford: Oxford University Press.
Lycan, W. G. (1987), *Consciousness*, Cambridge, MA: MIT Press.
Marmodoro, A. (2013), "Aristotelian Hylomorphism without Reconditioning", in *Philosophical Inquiry*: 36, 5–22.
Oderberg, D. S. (2007), *Real Essentialism*, New York: Routledge.
Oderberg, D. S. (2014), "Is Form Structure?", in *Neo-Aristotelian Perspectives in Metaphysics*, edited by D. Novotný and L. Novak, New York: Routledge, 164–180.
Olson, E. T. (1996), *The Human Animal*, Oxford: Oxford University Press.
Plantinga, A. (1979), *The Nature of Necessity*, Oxford: Oxford University Press.
Quine, W. V. O. (1948), "On What There Is", in *From a Logical Point of View*, (1953), edited by W. V. Quine, Harvard University Press, 1–19.
Rea, M. C. (2011), "Hylomorphism Reconditioned", in *Philosophical Perspectives*: 25, 341–358.
Schaffer, J. (2009), "On What Grounds What", in *Metametaphysics*: *New Essays on the Foundations of Ontology*, edited by D. Chalmers, D. Manley and R. Wasserman, Oxford: Oxford University Press, 347–383.
Sider, T. (1993), "Van Inwagen and the Possibility of Gunk", in *Analysis*: 53, 285–289.
Sider, T. (2012), *Writing the Book of the World*, Oxford: Oxford University Press.
Van Inwagen, P. (1981), "The Doctrine of Arbitrary Undetached Parts", in *Pacific Philosophical Quarterly*: 62, 123–127.
Van Inwagen, P. (1990), *Material Beings*, Ithaca, NY: Cornell University Press.
Wiggins, D. (1968), "On Being in the Same Place at the Same Time", in *Philosophical Review*: 77, 90–95.
Worrall, J. (1989), "Structural realism: The Best of Both Worlds?", in *Dialectica*: 43, 99–124.
Young, J. Z. (1971), *An Introduction to the Study of Man*, Oxford: Clarendon Press.
Zimmerman, D. (2003), "Material People", in *The Oxford Handbook of Metaphysics*, edited by M. J. Loux and D. Zimmerman, Oxford: Oxford University Press, 491–526

Alfredo Tomasetta

Remnant-Persons: A Commonsense Defence of Animalism

Abstract: Peter van Inwagen is among the main contemporary proponents of animalism, the view in personal ontology according to which we are identical with human animals. This paper focuses on a major objection to animalism put forward by Mark Johnston: the remnant-person problem. Your head detached from the rest of your body seems to be a person; now, if s/he is not you, then one can bring a person into being simply by removing tissue from something, and this is absurd. If, on the other hand, s/he is you, then animalism is false. Van Inwagen can offer a solution to this puzzle by denying the assumption that a detached head is a person. I agree, but I claim that this can be done without assuming his well-known, and controversial, mereological theses.

Peter van Inwagen (see Van Inwagen (1990)) is a prominent advocate of animalism, the view in personal ontology according to which we, human persons, are identical with human animals.[1] While relatively neglected in the past, at least in the analytic tradition, animalism has nowadays gained centre-stage in the metaphysical debates concerning our basic nature and a number of arguments have been offered in favour and against the doctrine.

This paper focuses on one major objection to the animalist thesis, namely 'the remnant-person problem' put forward by Mark Johnston (2007, 2017).

While the literature has explored different strategies to counter Johnston's challenge,[2] I think a very natural, and plausible response has been curiously overlooked. The aim of this paper is to present this response.

The plan is as follows. Section 1 briefly sketches the animalist view of human persons. Section 2 presents the remnant-person problem. Section 3 first considers a van Inwagen-inspired answer to the problem, and then presents my own response to it. Section 4, finally, addresses a possible reply to my proposed solution to Johnston's problem.

1 Blatti (2016) and Olson (2007a), pp. 23–47, are very good presentations of the animalist theory. Besides Van Inwagen (1990), other notable contemporary animalists are De Grazia (2005), Olson (1997, 2003), Merricks (2001) and Snowdon (1990, 2014).

2 On these strategies see, especially, Olson (2017).

https://doi.org/10.1515/9783110664812-015

1 Animalism

Animalism says that human persons – that is beings like you and me – are identical with biological organisms belonging to the species *Homo sapiens*. Or, to put it briefly, animalism says that we are identical with human animals.

Peter van Inwagen, for example, is identical to a certain human animal, or organism, the one we ordinarily call 'his organism' or 'his body'.[3] Donald Trump is identical to another organism, Trump's body, and every one of us is identical with a human animal, the one ordinarily called our body.

So animalism is not the widely shared idea that a human person is *intimately related* to a human animal. The thesis is quite straightforwardly – and more controversially – that a human person is *numerically identical* with a human animal.

One thing to notice is that animalism does not imply that every person is identical with a human animal: there may well be *divine persons*, and they, surely, are not human organisms. Van Inwagen, an animalist and a Christian, thinks precisely this way. Animalism just says that *human* persons are identical with human animals.

So what, according to animalism, makes us – and, say, divine persons – *persons*?

As Johnston (2007), p. 46, notices, animalists are happy to use John Locke's famous characterization of a person: "A thinking intelligent being, that has reason and reflection, that can consider itself as itself, the same thinking thing, at different times and places". The idea is, then, that a person is something characterised by some sophisticated mental features: and if God, an immaterial substance, or van Inwagen, a human animal, have these features, then God or van Inwagen are persons – a divine and a human person, respectively.

This idea, notice, allows that something that is now a human person may continue to exist without being a person. Right now, the animalist says, I am identical to a human animal – call it A – and I am a human person. Yet animal A might well exist, in the future, in a persistent vegetative state. If this becomes the case (and I hope not!), then I – that is, organism A – will exist without being a person.

So, according to animalism, human persons are not always persons. Are they always animals? Here I will simply assume that they are. I will assume, that is, that "no one of us can cease to be an animal without thereby ceasing to be" (Johnston (2017), p. 89). While being a somewhat controversial point, in the present

3 Subtle differences can be, and have been, drawn between "organism" and "body". In the present context, however, there is no problem in equating the two terms.

dialectical context this move is an innocuous one. This is because Johnston too subscribes to this assumption in presenting his remnant-person problem for animalism: I have no problem conceding this.

2 The remnant-person problem

This section offers a reconstruction of the remnant-person problem for animalism which is based on Johnston (2007), pp. 44–46, and Johnston (2017), pp. 112–113. While Johnston himself has presented the problem in slightly different ways, I think that what follows preserves all the essential points in his line of reasoning.

Let us start by considering the following scenario. Mr. Brown's whole head is removed and then his torso is destroyed. When Brown's body was intact, his brain secured the kind of mentality sufficient for personhood, and when only Brown's head survives, the brain still functions in the relevant way for some time, so that it continues to secure the kind of mentality sufficient for personhood. And so Brown's head constitutes a person in John Locke's sense, a *remnant-person*. Let us call this person "*H*" (for "head").

Now: either *H* is identical with Brown or s/he is different from him. Suppose the latter. Then either 1) *H* was there all along, a distinct person from Brown, or 2) *H* came into being after removing tissue from Brown and then destroying it. If one subscribes to 1), then one ends up admitting too many persons, and this is surely to be avoided. But 2) cannot be reasonably held, either. This is because the following principle is eminently reasonable – indeed, as Johnston (2007), p. 47, suggests, it "organizes some of our thinking about persons and physical reality":[4]

> (No Creation)
> You do not cause a person to come into being by removing and then destroying tissue.[5]

Given that (No Creation) is quite certainly true, it follows that, contra 2), *H* cannot be a newly created person.

So one reaches the conclusion that *H* must be identical with Brown. Now, given that, quite plausibly, *H* is not an animal, this conclusion means that Brown,

4 See also Johnston (2017), p. 111.

5 For present purposes, there is no need to consider the actual principle Johnston (2017), p. 111, endorses: "You don't cause a person to come into being by removing, disabling or destroying tissue, unless this positively causally impacts the neural basis of a capacity for reflective mental life, for example by removing a suppressor of that capacity".

a human person, has ceased to be a human animal without thereby ceasing to be. But, according to animalism, no human person can cease to be an animal without thereby ceasing to be. Hence animalism is false.[6]

3 A commonsense reply

3.1 Separated heads do not exist: The van Inwagen-inspired reply

In order to discuss some possible replies to the remnant-person problem, the following semi-formal presentation of the problem itself may be helpful.

1)	H is a person.	Premise
2)	H is not an animal.	Premise
3)	If H is a person, then either H = Brown or $H \neq$ Brown.	Premise
4)	Either H = Brown or $H \neq$ Brown.	From 1 and 3
5)	If $H \neq$ Brown , then either H coexisted with Brown or (No Creation) is false.	Premise
6)	H did not coexist with Brown.	Premise
7)	(No Creation) is true.	Premise
8)	It is not the case that ($H \neq$ Brown).	From 5, 6 and 7
9)	H = Brown.	From 4 and 8
10)	If (H = Brown and H is not an animal), then animalism is false.	Premise
11)	Animalism is false.	From 2, 9 and 10

Having clarified the assumptions of the argument, let us now consider what its more questionable premises are.

6 In reconstructing Johnston's reasoning, I have deliberately avoided his characteristic oscillation between "the head" on the one hand, and "the head and its artificial sustainers" on the other hand. Johnston, in fact, often talks about a head that is removed *and then kept alive and functioning by highly artificial means*. Yet, if H really were the-head-and-its-artificial-sustainers, then the animalist would have at her disposal an easy way out. She could simply say that H is a newly created person, and that this is perfectly compatible with the truth of (No Creation), because the remnant-person would have been created not just by removing and then destroying tissue, but also by connecting a human head to artificial sustainers. So, if Johnston's argument is to be at least a *prima facie* threat to animalism, then H has to be just a separated head. Indeed, Johnston (2017), while occasionally referring to the-head-and-its-artificial-sustainers, makes it clear that what is really at issue is just the head, pure and simple: the remnant-person is there even *before* artificial sustainers are involved.

Given that premises 3 and 5 are not controversial, that premises 6 and 7 are very reasonable, and that premise 10 is both reasonable and assumed as true in the present discussion, one is left with premises 1 and 2.

Now, someone might entertain the view that a separated head is indeed an animal, even though an animal in a radically mutilated condition.[7] And if this were true, one could block Johnston's argument by denying premise 2. Yet, to my mind, this is not a promising strategy: the idea that a separated head is an animal is decidedly not very appealing, and I will set it apart without further discussion.[8]

So the only remaining premise is premise 1: *H* is a person. And this is a thesis that may be challenged by recurring to well-known mereological ideas by van Inwagen.

According to Van Inwagen (1990) the only composite entities are the living ones: the spatio-temporal world features just two kinds of entities, simple particles and living organisms. And this means, for example, that there are no tables, chairs or mountains. There are, to be sure, particles 'arranged table-wise', 'chair-wise' or 'mountain-wise', but they do not compose anything. And there are also particles arranged "separated-head-wise" but they do not compose a head.

If one endorses this mereological doctrine, one can say that in the Brown story there is no separated head. The only material things one obtains after the removal of tissue from Brown are particles arranged head-wise. Hence, and quite simply, van Inwagen is able to say that *H* does not exist at all and, if so, *H* certainly is not a person and premise 1 is false.

Now, while I am quite sympathetic towards van Inwagen's mereological stance, recurring to a controversial quasi-nihilistic mereoleogy to address Johnston's challenge may well appear to be too drastic and onerous a move – so drastic and onerous that many neutral observers may remain unconvinced. So if there were a more down-to-earth answer to Johnston's concern, that answer would certainly be preferable, at least for dialectical purposes.

And a more down-to-earth defence of animalism also seems quite appropriate to the animalist doctrine itself, a doctrine that sounds like a very down-to-earth view of our nature.

What I am going to offer in what follows then is a way to deny premise 1 that is not based on van Inwagen's mereological quasi-nihilism, and that should sound somewhat mundane.

---

7 Lim (2011) argues for this thesis, and Johnston (2017), p. 122, says that in some places van Inwagen takes this view as well.

8 If one wanted more arguments, Johnston (2017), pp. 122–125, offers compelling reasons against the thesis that a separated head is an animal.

3.2 Separated heads do exist, but they are not persons: A commonsense reply

Let us concede to Johnston – and to common sense – that separated heads do in fact exist. But even though Brown's separated head exists, is this head really a person? Johnston (2017), p. 113, says it is, because when only Brown's head survives, the brain still functions in the relevant way for some time, so that it continues to secure the kind of mentality sufficient for personhood.

But this crucial assumption of Johnston's is far from being an obvious one. While it may well be that, after the tissue removal, the brain still functions in some way, the idea that it functions in a way apt to secure the mental life *of a person* seems to me quite controversial to say the least.

A person is something endowed with *very* sophisticated mental features, and one cannot be so sure that the brain of a separated head is really able to secure such a sophisticated mental life.

This observation alone, I think, casts some doubt on premise 1. But it is not just that premise 1 is far from being obvious; I also think that one can argue that it is most probably false.

To see this, consider again Locke's characterization of a person – which both animalists and Johnston subscribe to: "A thinking intelligent being, that has reason and reflection, that can consider itself as itself, the same thinking thing, at different times and places".

Is *H*, the separated head, a person in the Lockean sense? As I said, it is reasonable to say that the brain inside *H* still functions in some way, and so one may concede that *H* is indeed a thinking being – even though *H*'s thoughts will be of a peculiarly vanishing kind.

But can *H* consider itself as itself in different times and places? One might think it cannot, arguing that *H* has only momentary thoughts so that it cannot have thoughts at different times. This however is not a very convincing idea.

First of all, while *H* is indeed very short-lived, its life – including its mental life – plausibly extends for more that just one moment. And secondly, even conceding that *H* is indeed a momentary thinker, one may say, following Olson (2007b), p. 42, that it *can* consider itself as itself in different times and places: just as a poor man might consider himself rich, a momentary thinker might mistakenly consider itself as existing at different times and places.

So *H* is plausibly a thinking being, and, perhaps, one cannot exclude that *H* can consider itself as itself in different times and places. Should one then think that *H* is after all a person? I do not think so.

Consider the three further features a Lockean person must have, namely intelligence, reason, and reflection. And imagine someone, holding *H* – the separated

head – by the hair, saying that *H* is intelligent and reflective, and indeed a true reasoner.

What would you think about these statements? I guess you would be very sceptical and even incredulous. Indeed, I believe most of us would consider the idea that *H* is an intelligent and reflective reasoner as nothing more than a grotesque joke. And this, I submit, is a well-founded reaction.

According to the on-line Collins Dictionary[9], if something is intelligent, it has the ability to learn things quickly and well. Is *H* able to learn things quickly and well? Of course not, and indeed it seems plausible to say that *H* is not able to learn anything at all. And if so, *H* cannot sensibly be considered an intelligent being.

Let us then consider reflection. As the Collins Dictionary says, if one is reflective, one is thinking deeply about something.[10] Now, "thinking deeply" is obviously a vague expression, just as "being bald": baldness and depth come in degree and they have borderline cases. Yet, as there are clear cases of bald people, there are also clear cases of thinking that is definitely not 'deep'. Bruce Willis is clearly bald, everyone agrees. And I feel quite certain that an unbiased judge would also say that a separated head is a clear case of a thing that does not think deeply about anything.

So *H*, very plausibly, is neither intelligent nor reflective. And it has no reason, either, I would add.

Having reason, the Collins Dictionary recalls us, comprises the ability to make sensible judgements and to decide that something is true after thinking carefully about all the facts.[11] But just as it lacks the ability to learn things quickly and well, a separated head is simply not able to do these things, and hence it does not have reason.[12]

But if *H*, the separated head, has no intelligence, no reflection and no reason – as it seems very reasonable to say –, then *H* itself is not a person, and premise 1 of Johnston's argument can be rejected.

9 https://www.collinsdictionary.com/dictionary/english/intelligent

10 https://www.collinsdictionary.com/dictionary/english/reflective

11 https://www.collinsdictionary.com/dictionary/english/reason

12 Here I am basing my case on the quite obvious idea that *H* does not have the abilities at issue, without proposing an explanation of this fact. But, whatever way this explanation might go, it will not say that *H* lacks the relevant abilities *because H itself is very short-lived*. To see this, imagine God creating a perfect duplicate of Obama living for just one moment before the Creator annihilates it. The duplicate, being a momentary existent, is very short-lived, and yet it *has* the ability to make sensible judgements, to decide that something is true after thinking carefully, and to learn things quickly and well. It is just that it is not offered the possibility of exercising the abilities it possesses.

4 Defusing 'the remnant-thinker problem'

Let us suppose that one accepts what I said about *H* in the previous section: while being a thinking object, *H* itself is not a person, and so the remnant-*person* problem for animalism cannot be formulated.

Still, one may observe, if one concedes that *H* is a thinking being, then a remnant-*thinker* problem – which exactly parallels Johnston's original argument – can be put forward.

Here is the new problem.

> When Brown's whole head is removed, and his torso destroyed, one obtains *H*, a thinking being. Either *H* is identical with Brown or *H* is different from him. Suppose the latter. Then either 1) *H* was there all along, a distinct thinker from Brown, or 2) *H* came into being after removing tissue from Brown and then destroying it. If one subscribes to 1), then one ends up admitting too many thinkers, and this is surely to be avoided. But 2) too cannot be reasonably held. This is because the following principle is eminently reasonable:
>
> (No Creation*)
> You do not cause a thinker to come into being by removing and then destroying tissue.
>
> Given that (No Creation*) is quite certainly true, it follows that, *contra* 2), *H* cannot be a newly created thinker.
> So one reaches the conclusion that *H* must be identical to Brown. Now, given that, quite plausibly, *H* is not an animal, this conclusion means that Brown, a human person, has ceased to be a human animal without thereby ceasing to be. But, according to animalism, no human person can cease to be an animal without thereby ceasing to be. Hence animalism is false.

Is this an effective argument against animalism? I think it is not, and this is because (No Creation*) is very far from being "quite certainly true". Let me elaborate on this.

A person is a very sophisticated kind of thinker, endowed with very complex mental abilities that are able to sustain a high-level kind of mental life.

A mere thinker, on the contrary, can be a very simple being with respect to its mental life. Molluscs, flies and even viruses, for example, may well count as minded, thinking beings, but their mental capabilities are, most plausibly, rather elementary ones. And separated heads too are thinkers, while not having very refined mental abilities.

Now, I believe that the difference between persons and mere thinkers grounds a difference in the intuitive statuses of such principles as (No Creation) and (No Creation*).

While (No Creation) is, indeed, quite certainly true, this is not the case for (No Creation*).

The possibility of creating a *very* complex mental being such as a person by simply removing and destroying tissue sounds a rather weird idea.

But when it comes to mere thinkers, things are far less clear. A thinker can be a very elementary minded being and, I submit, there is no clear intuition that a very elementary minded being cannot be created by simply removing and destroying tissue.

My point, therefore, is not that (No Creation*) is certainly false, but that it is a quite controversial principle. And a very controversial thesis cannot be legitimately invoked as obviously or even plausibly true in deciding between rival philosophical theories.

If so, one cannot legitimately use the remnant-thinker problem in order to decide whether animalism is true or false. And this is because this argument is crucially based on a controversial principle that cannot be safely employed in the resolution of philosophical controversies.

A slightly different, and perhaps helpful, way of putting my point, is the following.

A standard method to try and adjudicate a philosophical controversy is to compare how rival theories behave with respect to some clearly true principles. If theory *T* does vindicate a larger number of these principles than its rivals do, this is one reason to prefer *T* over its contenders.

And once 'a winning theory' has emerged, one uses this theory to settle more controversial principles that, because of their controversial status, cannot be used to determine what the right philosophical theory is.

My point, then, can be expressed by saying that (No Creation) belongs to the clearly true principles that could be used to decide between rival theories, while (No Creation*) is among those controversial principles that are to be settled after a winning theory has emerged. And if this is so, (No Creation*) has to be decided by a theory, and one cannot use it as a crucial premise in an argument designed to argue for or against a philosophical thesis. So, again, one cannot legitimately use the remnant-thinker problem in order to decide whether animalism is true or false.

Yet, this does not mean that the remnant-thinker problem is completely useless. Suppose that after intense discussion, philosophers agree that animalism is indeed the winning theory as regard to our metaphysical nature. And having a winner, we want to use it to decide the status of some controversial principles such as (No Creation*). So what does animalism say about (No Creation*)? On the

face of it, the winning theory seems silent on the issue. And here is where the remnant-thinker problem shows its usefulness.[13]

An animalist cannot say that Brown is identical to *H*: no human person can cease to be an animal without thereby ceasing to be. So *H* is different from Brown. But *H* cannot have co-existed with Brown – there would be too many thinkers –, and so *H* came into being after removing tissue from Brown and then destroying it. Hence, according to animalism (No Creation*) is indeed a false principle – and the remnant-thinker problem has shown its usefulness in letting us know what animalism has to say on this intuitively controversial thesis.

5 Conclusion

In this paper I have argued that the remnant-person problem for animalism put forward by Johnston can be answered in a straightforward way: the alleged remnant-person featuring in the problem is not after all a person.

I also argued that the parallel 'problem of the remnant-thinker' one may raise is not a real threat for animalism.

In both cases I offered what I take to be rather commonsensical considerations, and so I think my defence of animalism appropriately qualifies as a commonsense reply to Johnston's challenge.

In this respect my strategy differs from the response to Johnston one might give using van Inwagen revisionist mereology.[14] As I said, I am quite sympathetic with this radical mereoleogical doctrine, but those friends of animalism who do not feel at home with it may find the response I have offered in this paper a proposal worth their consideration.

13 Of course, the remnant-thinker problem may be useful in many other ways I am not considering here.

14 I know van Inwagen would say his mereological ideas are not revisionist with respect to common sense, because he simply denies "that there is any such thing as the body of doctrine that philosophers call common sense" (Van Inwagen (1990), p. 103). I disagree, but this is an issue for another occasion.

Bibliography

Blatti, S. (2016), "Animalism", in *The Stanford Encyclopedia of Philosophy* (Winter Edition), edited by E. N. Zalta. URL = <https//plato.stanford.edu/archives/win2016/entries/animalism/>

De Grazia, D. (2005), *Human Identity and Bioethics*, Cambridge: Cambridge University Press.

Johnston, M. (2007), "'Human Beings' Revisited: My Body is not an Animal", in *Oxford Studies in Metaphysics*, Vol. 3, edited by D. W. Zimmerman, Oxford: Oxford University Press, 33–74.

Johnston, M. (2017), "Remnant Persons. Animalism's Undoing", in *Animalism*, edited by S. Blatti and P. F. Snowdon, Oxford: Oxford University Press, 89–127.

Lim, J. (2011), *Bodies and Persons*, Doctoral Dissertation, University of Virginia.

Merricks, T. (2001), *Objects and Persons*, Oxford: Oxford University Press.

Olson, E. T. (1997), *The Human Animal*, New York: Oxford University Press.

Olson, E. T. (2003), "An Argument for Animalism", in *Personal Identity*, edited by R. Martin and J. Barresi, Oxford: Blackweel, 318–334.

Olson, E. T. (2007a), *What Are We?*, New York: Oxford University Press.

Olson, E. T. (2007b), "What are We?", in *Journal of Consciousness Studies*: 14(5-6), 37–55.

Olson, E. T. (2017), "The Remnant-Person Problem", in *Animalism*, edited by S. Blatti and P. F. Snowdon, Oxford: Oxford University Press, 145–161.

Snowdon, P. F. (1990), "Persons, Animal, and Ourselves", in *The Person and The Human Mind*, edited by C. Hill, Oxford: Clarendon Press, 83–107.

Snowdon, P. F. (2014), *Persons, Animals, Ourselves*, Oxford: Oxford University Press.

Van Inwagen, P. (1990), *Material Beings*, Ithaca, NY: Cornell University Press.

Part V: Abstract Beings, Nominalism, and Infinity

Francesco Orilia

Van Inwagen's Approach to Relations and the Theory of O-Roles

Abstract: In "Names for Relations," Van Inwagen presents a semantic problem about names for non-symmetric relations, based on the presupposition that there are converse relations. This paper first provides an analysis of Van Inwagen's problem and then addresses it by appealing to o-roles such as 'agent,' 'patient,' 'theme,' instrument, etc., and by distinguishing sparse and abundant levels of properties and relations. At the sparse level, o-roles are appealed to in order to tackle the problem of relational order and reject converse relations. At the abundant level, converse relations are recovered in the light of the sequential character of thinking, speaking and writing, and Van Inwagen problem is solved by giving due consideration to the order in which relata are assigned o-roles and referred to.

1 Introduction

In his intriguing 2006 paper "Names for Relations" (*NR*, in short), Peter Van Inwagen presents a problem about names for non-symmetric relations and corresponding converses; the problem is somehow articulated into two distinct puzzles. These puzzles are semantic, but they ultimately address two interconnected ontological issues: whether there are converse relations and the problem of relational order, which deeply engaged Russell (1903, 1984). I shall discuss the two puzzles in §2, and the two ontological issues, in §3. Then, in §4, I shall present my view on these matters and tackle the naming problem raised by Van Inwagen. My approach is based on (i) allowing for both an abundant and a sparse realm of properties and relations, (ii) appealing to *onto-thematic-roles*, or, in short, *o-roles*, which are ontological counterparts of the linguist's thematic roles (agent, patient, theme, source, goal, instrument, etc.). As it will be clear from the examples that we shall discuss, almost all derived from *NR*, Van Inwagen endorses an abundant realm of properties, relations and propositions (*NR*, p. 456), within which converse relations are taken for granted. In contrast, I distinguish a sparse realm, in which o-roles allow us to dispense with converse relations, and an abundant realm, wherein there are converse relations. In the following, what is attributed to Van Inwagen, possibly with specific page or section references, is relative to *NR*.

https://doi.org/10.1515/9783110664812-016

2 Van Inwagen's puzzles

Van Inwagen argues that there are two puzzles regarding non-symmetric relations, which have no counterparts in the case of propositions, properties and symmetric relations: *the puzzle of names for non-symmetrical relations and the puzzle of relation sentences*, as they are labelled at the very end of *NR* (pp. 473–474). The puzzles are also cast, at the beginning of *NR* (p. 453), as a single problem, *the problem of finding a formal, systematic way of naming relations*, which then bifurcates into the two puzzles (discussed, respectively, in §2 and §3 of *NR*). The problem depends, as Van Inwagen tells us, on an assumption that he refuses to forego (p. 453). This is in fact a three-fold assumption, and so, we may say, there are really three assumptions in play, which are as follows (quoting from *NR*, except for the labels):

(A1) Every dyadic relation has at least one converse;
(A2) there are non-symmetrical dyadic relations;
(A3) no non-symmetrical dyadic relation is identical with any of its converses.

It is important to note that Van Inwagen takes for granted that we can *express* non-symmetric relations just as we can express propositions, properties and symmetric relations. We express propositions with sentences such as "the Taj Mahal is white" and we express properties and relations with predicative phrases such as "is white," "live in the same city," "loves," "is loved by," "is to the north of," "is to the south of." The pairs "loves"/ "is loved by" and "is to the north of"/"is to the south of" provide examples of converse relations: what we express with the second member of the pair is a converse of what we express with the first member of the pair (or vice versa). We may add, since it seems to be implicit in *NR*, that we can express properties and relations also with *open* sentences involving variables, such as "x is white," "x and y live in the same city," "x loves y," "x is loved by y," "x is to the north of y," "x is to the south of y" (It will be convenient to say that such open sentences have the form $S(x)$, $S(x, y)$, etc., depending on the variables that occur in them.) In sum, there is no problem in expressing non-symmetric relations and corresponding converses. But there are, as we shall see, problems in naming them, in designating or referring to them with *names* in a broad sense, i.e., expressions that can stand in subject position, more precisely, as we shall see, problems with *formal and systematic* names; I have said "names in a broad sense," because these formal and systematic names turn out to look, in at least some of the options considered by Van Inwagen, more like definite descriptions than names. Before moving to them, let us clarify, as regards (A1), that Van Inwagen does not take for granted that, for any relation, there is just one con-

verse (see *NR*, p. 474, note 4), since he thinks that, given a relation *R*, different co-extensive relations may count as converses of *R*. He admits however that we may single out in at least some cases a most salient converse of a given relation (though with some reservations due to his general concern about being really able to name non-symmetric relations; see p. 457). For example, 'being loved by' could be taken to be the most salient converse of 'loving,' which may however have other converses co-extensive with 'being loved by,' e.g., I surmise, 'being self-identical and loved by.'[1]

Let us now turn to formal and systematic names for propositions and properties. A formal and systematic name for a proposition is a term obtained from a sentence expressing that proposition by prefixing the locution "the proposition that;" for example, the sentence "the Taj Mahal is white" gives rise to the formal and systematic name "the proposition that the Taj Mahal is white" (p. 454). Analogously, a formal and systematic name for a property is a term obtained from an open sentence of the form $S(x)$, which can be taken to express the property in question, by prefixing the locution "the property of being an x such that" (p. 456). For example, the open sentence "x is white" gives rise to the formal and systematic name "the property of being an x such that x is white."[2] These names are governed by truth conditions of this sort: the proposition that P is true if and only if P; something, a, has the property of being an x such that $S(x)$ if and only if $S(a)$. For instance, the proposition that the Taj Mahal is white is true if and only if the Taj Mahal is white; the Taj Mahal has the property of being an x such that x is white is true of if and only if the Taj Mahal is white.[3]

A formal and systematic name for a dyadic relation should similarly be, suggests Van Inwagen, a term obtained from an open sentence of the form $S(x, y)$,

1 In an effort to improve readability I resort here and elsewhere in the following to the convention of using single quotes to designate relations (and possibly other abstract entities such as o-roles or positions, which will come to the fore below). I shall avoid however these single quotes when readability does not seem to me in danger. To designate linguistic expressions I use instead double quotes. Note that Van Inwagen uses both single quotes and double quotes to refer to linguistic expressions, although occasionally he also uses double quotes to single out relations.

2 As so constructed, these "names" look more like definite descriptions than names. But I think that we could without expressive loss have "that P" instead of "the proposition that P," and "being an x such that $S(x)$," instead of "the property of being an x such that $S(x)$."

3 Although Van Inwagen is not explicit about these truth conditions, that he takes them for granted can be inferred from the "semantics" provided in §3, despite the fact that he uses there a rather idiosyncratic terminology based on the locution "enter into." There are other terminological alternatives that could be used instead of "a *has* the property of being an x such that $S(x)$:" "the property of being an x such that $P(x)$ *holds* of a," "*a instantiates* (or *exemplifies*) the property of being an x such that $S(x)$," "the property of being an x such that $S(x)$ is *true* of a."

which can be taken to express the relation in question, by prefixing a locution analogous to "the property of being an *x* such that," except that it has "relation" instead of "property" and contains two variables instead of only one, the same variables that we find in the open sentence; and similarly for relations of further adicity (for simplicity's sake, I shall restrict my attention as much as possible to the dyadic case, since the issues we shall discuss already arise at this level and there will usually be no need to complicate the discussion by considering the other cases). In essence, the first puzzle, the puzzle of names for non-symmetrical relations, amounts, I would say, to this: we do not succeed in obtaining in this way names for non-symmetric relations, whose existence is granted by (A1), because a candidate name *N* of this sort suffers from semantic indeterminacy: it might be taken to name either a certain relation, say *R*, or a corresponding converse, say *R** (which, in view of (A2) and (A3) exists and is different from *R*), and there is no way to tell whether *N* names *R* or *R**, or, we may even say, there is no truth of the matter about this. Let us see why.

As a prefix analogous to "the property of being an *x* such that," Van Inwagen first proposes the locution "the relation that holds between *x* and *y* if and only if *S*(*x*, *y*)." This proposal is perplexing, because on analogy with names for properties, one would expect this: "the relation of being *x* and *y* such that *S*(*x*, *y*)." But let us sidestep this for the time being (I shall recover this option in §4). Following the proposal, suggests Van Inwagen, from the open sentence "*x* and *y* live in the same city," we can obtain the name "the relation that holds between *x* and *y* if and only if *x* and *y* live in the same city," which refers to a certain symmetric relation, expressed by "*x* and *y* live in the same city," and which we could more informally call "living in the same city;" the one that holds between, say, Barak Obama and Michelle Obama, or Michelle Obama and Barak Obama, no matter the order with which we refer to the former presidential couple. And things go similarly for other symmetric relations. Analogously, we may at least prima facie think, from the open sentences "*x* loves *y*" and "*x* is to the north of *y*," one may derive the names "the relation that holds between *x* and *y* if and only if *x* loves *y*," and "the relation that holds between *x* and *y* if and only if *x* is to the north of *y*," where the former refers to the relation expressed by "*x* loves *y*," which we may more informally call "loving," and the latter to the relation expressed by "*x* is to the north of *y*," which we could more informally call "being to the north of." But things do not go as smoothly now. For example, "the relation that holds between *x* and *y* if and only if *x* loves *y*" may be taken to refer to either 'loving' or 'being loved by;' and "the relation that holds between *x* and *y* if and only if *x* is to the north of *y*" may be taken to refer to either 'being to the north of,' or to 'being to the south of.' For both 'loving' and 'is loved by' hold between Harry and Sally if and only if Harry loves Sally; and both 'being to the north of' and 'being to the south of' hold between

Denmark and Italy if and only if Denmark is to the north of Italy. The problem, explains Van Inwagen, is that these candidate formal and systematic names, because of the indifference to order of the "and" in their subpart "… holds between … and …," fail to specify the order in which relata enter a relation (p. 459):

> The problem, of course, is created by the fact that a predicate like 'holds between Denmark and Italy' does not specify the order in which those two nations are to enter into the relation of which it is predicated: 'holds between Denmark and Italy' and 'holds between Italy and Denmark' are synonymous.

As we already saw, in order to convey that two objects are related by a relation, R, Van Inwagen says that R holds between them. In this quotation he also says that they *enter into* R. At other places, he says that they *stand in* R, or, as I have done myself at the beginning of this sentence, that they *are related* by R. Whatever the locution, what is emerging in the passage above is the presupposition that, when a non-symmetric relation R holds between two objects, it does so in one or the other of two possible *orders* (p. 459, p. 465), say order O_1 and order O_2, and these orders are such that, if R holds between these two objects in order O_1, then a converse of R holds between the same objects in the other order O_2, and vice versa; that is, if R holds between the two objects in question in order O_2, then the converse of R holds between the same objects in the other order O_1. Alternatively, in other parts of the paper (e.g., pp. 473–474) he speaks of different ways in which a non-symmetric relation, R, holds between two objects, say way W_1 and way W_2, which are such that if R holds between two objects in way W_1, then a converse of R holds between the same objects in the other way W_2. To anticipate, I myself think that it is appropriate to say that a non-symmetric relation is jointly exemplified in two different ways by two relata, or in Van Inwagen terminology, that two relata enter a non-symmetric relation in two different ways. But these ways, which I shall identify with o-roles, should be kept distinct from the orders that Van Inwagen has in mind here, which are really *sequential* orders. Let us then avoid the "way" terminology at this juncture; I shall use it in my own manner in § 4.

Let us go back to the problem at issue. Concerning 'being to the north of,' we can make two hypotheses: either (i) it holds of x and y in order O_1 when x is to the north of y, in which case a converse, say 'being to the south of,' holds of x and y in order O_2; or (ii) it holds of x and y in order O_2 when x is to the north of y, in which case 'being to the south of' holds of x and y in order O_1. Now, the description "the relation that holds between and x and y if and only if x is to the north of y" is improper (p. 459), because it may be understood either as (a) "the relation that holds between x and y in order O_1 if and only if x is to the north of y," or as (b) "the relation that holds between and x and y in order O_2 if

and only if x is to the north of y;" but in either case, it is undeterminate whether it refers to 'being to the north of,' or to its converse, 'being to the south of,' for this depends on whether we assume (i) or (ii). Consider first hypothesis (i). In this case, the relation that holds of x and y with order O_1 is 'being to the north of,' and consequently 'being to the south of' is the relation that holds of x and y in order O_2. Hence, given interpretation (a), the description refers to 'being to the north of,' and given interpretation (b), it refers to the converse 'being to the south of.' Consider now hypothesis (ii). Given interpretation (a), it refers to 'being to the south of,' and given interpretation (b), it refers to 'being to the north of.'

We may put matters as follows. Since "and" is indifferent to order, "the relation that holds between x and y if and only if x loves y" and "the relation that holds between y and x if and only if x loves y" are synonymous (p. 45). Hence, both of them name the same relation, if any, but we have a semantic indeterminacy, because we have no way to tell whether what they name is 'loving' or 'being loved.'

Van Inwagen then considers the option of obtaining a formal and systematic name by prefixing to the open sentence a locution that succeeds in specifying the order in which the relation in question holds between the relata. He then proposes these further candidates: "the relation that holds between x and y, in that order, if and only if $S(x, y)$" or "the relation that holds between x and y, in the order x first, y second, if and only if $S(x, y)$." But Van Inwagen is now skeptical about making sense of this specification of order, since he takes it as an invitation to really accept the idea that relata enter a relation in a sequential order, one first, the other second, and yet he doubts that we can really understand this. He expresses this skepticism with these words (p. 460):

> But what is it for a relation to hold between – for example – Italy and Denmark in the order "Denmark first, Italy second"? You may well ask.

This raises a perplexity, since in order to convey the puzzle at issue, as we have seen, Van Inwagen seems to take for granted that relata enter a relation in an order.

Let us now turn to the second puzzle, the puzzle of relation sentences. In logic we use expressions that seem to name relations, whether symmetric or not, in a formal and systematic way. They are the so called *lambda-abstracts*, obtained by prefixing the operator "λ," followed by n variables, to an open formula containing those variables, so as to obtain a term for an n-adic relation. Van Inwagen considers whether we could use the same device with open sentences of English so

as to get the formal and systematic names he is looking for.[4] Consider the dyadic case: given an open English sentence of the form $S(x, y)$, we obtain the formal and systematic name "$\lambda xy\ S(x, y)$." For example, we can have names such as "λxy x loves y" and "λxy x is to the north of y." Actually, Van Inwagen simplifies matters by proposing a notation that drastically cuts the number of lambda-abstracts that we can create. This is achieved by using boldface numerals instead of variables bound by the lambda operator and by ruling that one always resorts to the lowest possible numerals. To better signal the change of notation Van Inwagen also uses a different symbol, "R" (followed by an expression within parentheses), instead of "λ." Thus, for example, whereas in the more familiar notation we could write "λxx x loves x," "λyy y loves y," "λxy x loves y," "λyx x loves y," "λxy y loves x," "λyx y loves x," "λzz z loves z," in Van Inwagen's notation we can write only "R(1 loves 1)," "R(1 loves 2)," "R(2 loves 1)." It is indeed a good idea to cut down on possible symbols in this way, since we avoid notational variants without inducing real expressive limitations. But we can obtain the same goal by ruling that (i) the variables after "λ" should be in alphabetic order and (ii) the variables that get to be used are as lower as possible in the alphabetic order (starting as usual with "x"). Thus, we can only have "λxx x loves x," "λxy x loves y" and "λyx x loves y." With this in place, we can keep using the more familiar lambda notation.

Despite these niceties, Van Inwagen does not think that these *abstraction-operator names*, as he calls them, really provide a solution to the problem of finding formal and systematic names for non-symmetric relations. We may think that we succeed in grasping what the abstraction-operator names refer to in the light of the truth conditions (or "semantics," in Van Inwagen's terminology[5]) with which we are supposed to use such names (which are basically what logicians call principles of *lambda-conversion*). They are as follows:

> t_1 and t_2 enter into $\lambda xy\ S(x, y)$ if and only if $S(t_1, t_2)$;
> t_1 and t_2 and t_3 enter into $\lambda xyz\ S(x, y, z)$ if and only if $S(t_1, t_2, t_3)$;
> etc.

4 In logic lambda-abstracts are usually taken to be predicates, rather than names, but they are also used as names in the way Van Inwagen envisages, in particular in type-free property theories (see Orilia and Swoyer (2017), §8).

5 We neglect here details of this semantics that Van Inwagen inserts in order to eschew problems such as Russell's paradox and empty terms, as they are not relevant for what we are discussing here.

Here are some examples for the dyadic case:

> Tim and Tom enter into the relation λxy x and y live in the same city if and only if Tim and Tom live in the same city;
> Harry and Sally enter into λxy x loves y if and only if Harry loves Sally;
> Denmark and Italy enter into λxy x is to the north of y if and only if Denmark is to the north of Italy.

The left-hand sides of these biconditionals are called by Van Inwagen *relation sentences*, and the right-hand sides their *non-relational counterparts* (which explains why the second puzzle is called "the puzzle of relation sentences").

These truth conditions reveal that, for example, "Harry and Sally enter into λxy x loves y" is not to equivalent to "Sally and Henry enter into λxy x loves y." For the truth conditions tell us that the two sentences to which they are respectively equivalent are not equivalent to each other: "Harry and Sally enter into λxy x loves y" is equivalent to "Harry loves Sally," whereas "Sally and Henry enter into λxy x loves y" is equivalent to "Sally loves Henry." Nevertheless, claims Van Inwagen, these truth-conditions do not enable us to decide, e.g., which among "λxy x loves y" and "λxy y loves x" refers to 'loving' and which to 'being loved'; or which among "λxy x is to the north of y" and "λxy y is to the north of x" refers to 'being to the north of' and which to 'being to the south of' (pp. 467–468). And thus, even with lambda-abstracts we have semantic indeterminacy. In contrast, we may add, "λxy x and y live in the same city" refers precisely to one (symmetric) relation, the one expressed by "x and y live in the same city."

3 Relational order

It seems to me that in order to solve, or dissolve, these puzzles, we need to clarify the nature of relations and how they get to be jointly instantiated or exemplified by different objects, or, in Van Inwagen's terminology how objects enters into relations. Van Inwagen seems to infer from the existence of these puzzles that we do not really understand these things. But I think that some headway can be made by bringing to the fore the problem of relational order and then by addressing it by appealing to o-roles and to a distinction between the sparse and the abundant realm. As anticipated, rather than taking "way" and "order" as synonymous and expressing some idea that we do not clearly understand, I shall invoke two different notions. I shall speak of ways in which relations are exemplified, and I shall identify them with o-roles, which are invoked primarily at the sparse level. O-roles make relations, so to speak, "neutral" and thus devoid of converses. But at the

abundant level I shall take seriously the idea that the *sequential* order with which we designate relata matters somehow, thereby recovering converse relations.

The problem of relational order was so labelled by Hochberg (1987), who took it over from Russell; alternatively we may speak of the problem of differential application, in MacBride's terminology (MacBride (2016)). Whatever the name, it is the problem of explaining what accounts for the difference between two *distinct* facts or states of affairs such as Harry's loving Sally and Sally's loving Harry, which appear to involve the same relata, Harry and Sally in this case, and the same non-symmetric relation, an amatory relation in this case, which we may simply call *love* (thereby refraining from identifying it with 'loving' or 'being loved by;' perhaps it is neither). Following Hochberg, we may say that such states do not differ in their constituents, namely love, Harry and Sally, or, more conveniently, that they do not differ in their *canonical* constituents, so as not to rule out that they may have other constituents beside the relation in question and the relata. Yet they differ, and we may then say that they differ in *relational order*. But this is just a label; we need explain what this difference amounts to. That is, we want to explain what relational order is. It should be noted that, when symmetric relations are involved, there is no relational order. For example, if there is a state of affairs, s, consisting of Harry's being adjacent to Sally, there is no state of affairs with the same canonical constituents, 'adjacent,' Harry and Sally, and yet different from s. We may, by changing word order, speak of the state of affairs consisting of Sally's being adjacent to Harry, but by so speaking we appear to refer to the same old state of affairs, *s*. We may then say that a state of affairs such as this *lacks* relational order. Presumably, a good explanation of relational order should bring with it also an explanation of what *lack* of relational order amounts to.

In tackling these issues, Russell was guided by two contrasting intuitions. In *PoM* (1903), he took for granted, just like Van Inwagen, that non-symmetric relations have distinct converses: there exist, for instance, both 'loving' and the distinct converse 'being loved by,' both 'before' and the distinct converse 'after', etc. In Russell's view, they have different and opposite "directions" or "senses." With directions available we can account, Russell seems to have thought, for the difference in relational order between Harry's loving Sally and Sally's loving Harry (to still use, anachronistically, one of Van Inwagen's examples): one of them involves a relation with a certain direction, and the other a relation with the opposite direction, and so they differ, because they differ in one canonical constituent after all, namely the relation. In fact, this approach does not really solve the problem of relational order (see Orilia (2008)), but we need not dwell on this here. Let us rather note that, by distinguishing a non-symmetric relation and its converse, we seem to be driven to acknowledge that we have two distinct facts, when we seem to have just one fact. For example, if 'loving' and 'being loved by' are two distinct

relations, there seem to be two distinct facts, Harry's loving Sally and Sally's being loved by Harry, which differ in one constituent, namely 'loving' and 'being loved by.' In *ToK* (1984), Russell took for granted that this cannot be the case and accordingly rejected the idea that non-symmetric relations have distinct converses. Relations are now taken to be "neutral," devoid of any direction or sense. There is then only one neutral relation, love, and no 'loving' or 'being loved by' (at least at a fundamental level); or only one neutral relation, say *temporal succession*, and no 'before' or 'after.' In order to account for relational order we now need "positions." For example, with respect to the fact consisting of Harry's loving Sally, Harry has the 'lover' position and Sally the 'beloved' position. This fact differs in relational order from the fact consisting of Sally's loving Harry, since with respect to this further fact Harry has the 'beloved' position and Sally the 'lover' position. It should be noted that, since 'loving' and 'being loved by' as distinct relations are rejected, there is in this account no further fact which is Sally's being loved by Harry. This just is Harry's loving Sally. And similarly there is no further fact which is Harry's being loved by Sally.[6]

Following the lead of Russell, Hochberg (1987) explains relational order by appealing to *ordering relations* such as 'first,' 'second,' 'third,' etc. Just as in *ToK*, relations are neutral and thus there are no converse relations. And when a neutral relation is exemplified, thereby giving rise to a relational fact involving various relata, we can distinguish differing ordering relations that link the relata to the fact.[7] Thus, for example, Harry and Sally are linked to the fact consisting of Harry's loving Sally by the ordering relations 'first' and 'second,' respectively. In contrast, Harry and Sally are linked to the fact consisting of Sally's loving Harry by the ordering relations 'second' and 'first,'respectively. Thus, the two facts are different. It may seem that the Russell of *ToK* and Hochberg differ only verbally, but this is not the case. Hochberg's ordering relations are *inter-repeatable* in the sense that they can occur with different relations. Consider for example the relation of temporal succession. According to Hochberg, a succession fact consisting of the year 2000's being after the year 1980 would involve the very same ordering relations, 'first' and 'second,' which are also involved in an amatory fact such as Harry's loving Sally. In contrast, Russell's positions are not inter-repeatable: they can occur only with their own specific relations. Thus, a temporal succession fact

6 A view similar to this was independently developed and endorsed by Castañeda (1975), who traces it back to Plato and Leibniz (see Castañeda (1982)).

7 Actually, in Hochberg's account the relata are linked not to a fact, but to an "unordered complex" (see Orilia (2014) for a discussion of this). But this detail need not detain us here.

such as 2000's being after 1980 would involve two specific positions, say the 'predecessor' position and the 'successor' position.

The *PoM* account, in which we have, in agreement with Van Inwagen, converse relations, is also an account in which Russell acknowledges, just like Van Inwagen, an abundant realm of properties, relations and propositions (thus, what I said above as regards *PoM* had better be "translated" by understanding states of affairs as propositions, or by taking "fact" as synonymous with "true proposition" rather than with "state of affairs"). In contrast, in the *ToK* account, in which there are no converse relations, Russell seems inclined to accept something like a sparse realm of properties, relations and states of affairs (though he still labels them "propositions"). But I think that we should have it both ways and acknowledge *dualism* (see Orilia and Swoyer (2017), §5.4, and references therein), the view according which we have both an abundant realm and a sparse realm; at the *intensional* level and at the *truthmaker* level, respectively, we may say. For abundant properties, relations and propositions can account for natural language semantics and propositional attitudes, by providing semantic values for all predicates and sentences, with the appropriate fine-grained distinctions; whereas sparse properties, relations and states of affairs work as truthmakers, with the required coarse-grainedness. Consider for instance the true sentences "FO is both 1.82 meters and 70 kilos" and "FO is both 70 kilos and 1.82 meters." At the intensional level, we may say that the former expresses a proposition that attributes to FO the conjunctive property of being both 1.82 meters and 70 kilos, whereas the latter expresses another proposition, the one that attributes to FO this *other* conjunctive property: being both 70 kilos and 1.82 meters. However, at the truthmaker level, we need not invoke conjunctive properties, let alone a distinction between two conjunctive properties that differ only for the order of their conjuncts. All we need to say is that "FO is both 1.82 meters and 70 kilos" and "FO is both 70 kilos and 1.82 meters" are both made true by two states of affairs, namely the one consisting of FO's being 1.82 meters and the other one consisting of FO's being 70 kilos. If we grant this, we can then follow the Russell of *PoM* (and Van Inwagen) in admitting converse relations at the abundant level and the Russell of *ToK* in rejecting converse relations at the sparse level. With an important proviso in the latter case. Whereas Russell appeals to positions, I would appeal to o-roles, which are importantly different from positions and ordering relations in a manner that I shall now explain.

4 O-roles, the sparse realm and the abundant realm

As anticipated in the introduction, o-roles are ontological counterparts of linguistics' thematic roles such as agent, patient, source, theme, destination, etc. (I use for simplicity's sake the same terms for thematic roles and the corresponding o-roles, and rely on the context to distinguish them). I see o-roles as different ways in which objects jointly exemplify a neutral relation. Contrary to Russell's positions, they are inter-repeatable. In this they are analogous to Hochberg's ordering relations. But whereas the latter are, we may say, *promiscuously* inter-repeatable, in the sense that they can occur with all sorts of relations, o-roles are *moderately* inter-repeatable: they can occur with some different relations, but not with all sorts of relations. For example, 'agent' and 'patient' may occur in an analogous way with both love and hate, but have nothing to do with temporal succession. To illustrate, let us assume that Harry and Sally love each other and that Tim and Tom hate each other. From the point of view of the o-role theory, there are just two amatory states of affairs, Harry's loving Sally and Sally's loving Harry. They differ in relational order in that the former involves Harry with the 'agent' o-role and Sally with the 'patient' o-role, whereas the latter involves Sally with the 'agent' o-role and Harry with the 'patient' o-role. Similarly, there are just two "odiatory" states of affairs, Tom's hating Tim and Tim's hating Tom, in which the very same 'agent' and 'patient' o-roles somehow occur.

In all these approaches that we have seen, relational order can be viewed as the occurrence of different "order imposing" constraints on states of affairs, whether these constraints are positions, ordering relations or o-roles. In contrast, lack of relational order is the repeated occurrence of the very same constraint in a state of affairs. For example, there is relational order in Harry's loving Sally, because Harry and Sally are involved in it with two different positions, ordering relations or o-roles (namely, 'agent' and 'patient,' if we go for o-roles). In contrast, there is lack of relational order in Harry's being adjacent to Sally, because Harry and Sally are involved in it with the same position, ordering relation or o-role (namely, I would say, the 'theme' o-role, if we go for o-roles). The appeal to o-roles, however, is the best option, in my view, because it captures uniformities and differences in states of affairs, which otherwise are left unacknowledged; in this respect o-roles are pretty much like universals. For example, there is clearly a uniformity linking Harry's loving Sally and Tom's hating Tim, which can be captured by focusing on their both involving 'agent' and 'patient' o-roles, but there is no such uniformity in the year 1980's preceding the year 2000. On the other hand, there is perhaps a uniformity linking the latter to the number 5's being smaller

than the number 18, which can be captured by invoking in both cases the same pair of distinct o-roles, i.e., I am tempted to say, the 'source' and 'destination' o-roles.

Assuming then that there are o-roles, we may ask what they are and how they get involved into states of affairs. My proposal is that o-roles are *ways* in which relations are jointly instantiated. Thus, for example, the state of affairs consisting of Harry's loving Sally is the joint exemplification of the amatory relation by Harry in the way of being agent (*as* agent, in short) and by Sally in the way of being patient (*as* patient, in short). These ways are very generic properties that are exemplified by the relata inasmuch as they jointly exemplify the relation. To put it otherwise, when the relata jointly exemplify the relation, by the same token they also exemplify these very generic properties.

Once we appeal to o-roles, relations are neutral and do not split, so to speak, into pairs whose members are one the converse of the other. Van Inwagen's puzzles crucially rely on distinguishing relations and corresponding converses. These puzzles do not then arise at the sparse level in this approach. It is interesting to note, however, that we are able to take relations as neutral, by taking seriously the idea somehow hinted at by Van Inwagen, that there are distinct ways in which one relation is exemplified; these ways are identified with o-roles. Consider the two sentences

(1) Harry loves Sally;
(2) Sally is loved by Harry.

At the sparse level of truthmakers, there is one state of affairs involving the amatory relation that makes both of them true. By using "L" for this relation and obvious short labels for the o-roles and the proper names involved, we may formally represent this state of affairs as follows:

(1a) $L(Agt(h), Pat(s))$.

It should be noted that (1a) must be taken to imply that Harry is an agent and that Sally is a patient, that is: $Agt(h)$ and $Pat(s)$. It should also be noted that the order with which we refer to the relata is inessential. Thus, (1a) is equivalent to

(2a) $L(Pat(s), Agt(h))$.

A crucial issue in this approach is trying to understand which o-roles should ultimately be acknowledged. One could imagine two different strategies at this juncture. The *liberal* one is to freely postulate new o-roles whenever we encounter relations that seem to require o-roles other than the ones we are led to acknowledge

by relying on the thematic roles typically accepted in linguistics. The other, *restrained*, option is to limit as much as possible the o-roles we postulate, by relying on appropriate analyses of relations. We may very well focus on Van Inwagen's example, Denmark's being to the north of Italy, to illustrate the two options. Let us then consider the sentence:

(3) Denmark is to the north of Italy.

'Being to the north of' appears to be a non-symmetric relation and thus, following the liberal option, one may think that it requires two special o-roles, say N and S, associated to a neutral relation of spatial arrangement, say A. Thus, the truthmaker for (3) would be represented as follows:

(3*) $A(N(d), S(i))$.

Following the restrained option, before postulating new o-roles such as N and S, we first try to see whether the relation in question is analyzable. This is clearly possible, since x to be to the north of y really means something like this: that x lies in between y and the North pole, which we may designate with "n." In turn, that y is between x and z can be understood along these lines: b is in a spatial region delimited by x and y. As regards "in," linguists typically appeal to a 'theme' thematic role and we can correspondingly postulate a 'theme' o-role, which I shall abbreviate a "Th." For the delimiting objects, I think it is appropriate to appeal to a *boundary* o-role, Bnd in short, while presupposing a neutral relation of being connected by a region of space, call it R. By taking this line, we can represent the truthmaker for (3) as

(3a) $R(Th(d), Bnd(n), Bnd(i))$.

I think the restrained option is better. For one thing it is of course more economical. For another thing, it allows us to capture uniformities that would otherwise go unnoticed. For example, we may say that (3b) is analogous to the state of affairs consisting of Tom's being in Paris in that both involve the theme o-role. That Tom is in Paris could in fact be represented as follows (where "C" stands for a containment relation, "t" for Tom, "p" for Paris and "Loc" for the 'location' o-role): $C(Th(t), Loc(p))$.[8]

Let us now turn to the abundant realm. At this level we may want to say, as Van Inwagen does, that there are distinct relations expressed by "live in the same

8 See Orilia (2014) for additional details on this approach.

city," "loves," "is loved by," "is to the north of," "is to the south of," or, to pick another of "Van Inwagen's examples, "*x* loves *y* more than *z* loves *w*." (p. 468). From the fine-grained point of view of the abundant realm, we may very well say that even the writing, speaking or thinking order matters, for thinking, speaking and writing are sequential, and this is, I think, the key for distinguishing relations and their converses. After all, at this level we even differentiate between two conjunctive propositions such as the proposition that *P* & *Q* and the proposition that *Q* & *P*; we say that they are *two* propositions that entail each other, not that they are one and the same proposition. Similarly, we may want to say that (1) and (2) express different propositions, since they differ in the sequential order with which we designate the relata. And we may well capture this by taking (1a) and (2a) to represent these two different propositions, since after all (1a) and (2a) also differ in the order with which the relata are designated. Similarly, we may assume that (3) and

(4) Italy is to the south of Denmark

express two different propositions, the former perspicuously represented by (3a) and the latter by

(4a) $R(Bnd(i), Bnd(n), Th(d))$.[9]

Once we admit this, it is just a small step to acknowledge that, just as there are two different propositions represented by (1a) and (2a), there are also distinct relations linking Harry and Sally on the one hand, and Sally and Harry on the other hand, such that one is the converse of the other, namely $\lambda xy\ L(Agt(x), Pat(y))$, which can be naturally taken to be the relation expresses by "loves," i.e. 'loving', and $\lambda xy\ L(Pat(x), Agt(y))$, which can be naturally taken to be the relation expressed by "is loved by," i.e., 'is loved by.'

One may still wonder, however, what to make of the lambda-abstracts that we obtain from those we have just seen, by inverting the variables used in the open sentence within the scope of the λ-operator: "$\lambda xy\ L(Agt(y), Pat(x))$)" and "$\lambda xy\ L(Pat(y), Agt(x))$." Couldn't $\lambda xy\ L(Agt(y), Pat(x))$) be the converse of the relation expressed by "loves" or perhaps that relation itself? Do we still have

9 Strictly speaking I would rather say that (4a) perspicuosly represents a proposition that we could somehow, albeit clumsily, express in English by "Italy has, at North, Denmark." For perhaps "is to the south of" expresses a relation involving the South pole, rather than the North pole. But we may neglect these details for the sake of the point about sequentiality that I want to make here.

a semantic indeterminacy problem? Well, I think that we can continue cutting down on useless abstracts, following the lead of what Van Inwagen has already suggested. We may also rule that variables in the open sentence following the lambda variables must always be used in alphabetic order (except of course when we repeat a variable already used in order to signal a "reflexivization" as in the abstract "λxy x loves y more than x," which, to take note of the reflexivization, may be read as "being an x and a y, in that order, such that x loves y more than him/herself"). This rules out "λxy $L(Agt(y), Pat(x))$" ("λxy y loves x") and "λxy $L(Pat(y), Agt(x))$" ("λxy y is loved by x"). And thus we are left with only two candidate names (or less formal variations on the theme) for the relations expressed by "loves," and "is loved by," namely "λxy $L(Agt(x), Pat(y))$" ("λxy x loves y") and "λxy $L(Pat(x), Agt(y))$" ("λxy x is loved by y"). And o-roles and the sequential order can guide us in deciding which is which. We use "loves" with a term designating an agent in first position and a term designating a patient in second position, and it is the other way around with "is loved by." To put it otherwise, "$L(Agt(x), Pat(y))$" and "$L(Pat(x), Agt(y))$" are just ways of expressing – by explicitly invoking o-roles – what we express with "x loves y" and "x is loved by y," respectively. Thus, "λxy $L(Agt(x), Pat(y))$," with "Agt" in first position and "Pat" in second position is the obvious candidate for naming 'loving,' and "λxy $L(Pat(x), Agt(y))$," with "Pat" in first position and "Agt" in second position is the obvious candidate for naming 'is loved by.' And of course we can say similar things about 'being to the north of' and 'being to the south of,' and analogous pairs of relations and corresponding converses. And thus we have, I submit, a solution to the second puzzle posed by Van Inwagen, which regards the interpretation of the lambda-abstracts that are meant to stand for non-symmetric relations.

Moreover, by using "in that order" to appropriately emphasize that the sequential order in which we write the variables is semantically relevant, lambda-abstracts can be understood in English by resorting to a locution of this kind: "the relation of being an x and a y, in that order, such that $S(x, y)$."[10] But we could also avoid the use of variables and of "in that order," thereby eschewing Van Inwagen's misgivings about this locution, by saying "the relation of being a first individual and a second individual such that S (the first individual, the second individual). For example, we could say: "the relation of being a first individual and a second individual such that the first individual loves the second individual," corresponding to "λxy $L(Agt(x), Pat(y))$," and "the relation of being a first individual and a second individual such that the first individual is loved by the second individual,"

10 Or, more simply, "being an x and a y, in that order, such that $S(x, y)$."

corresponding to "$\lambda xy\ L(Pat(x), Agt(y))$." We thus have two formal and systematic names in English for 'loving' and 'being loved by,' of the sort that Van Inwagen is looking for. And similarly we can have names of this sort for 'being to the north of' and 'being to the south of,' etc. In sum, we also have a solution to the first puzzle, which regards finding formal and systematic names in English for non-symmetric relations.

It should be noted that no new o-roles are needed to attribute these relations of the abundant realm to objects; the appropriate o-roles are already inside the relations, so to speak, as it surfaces from the truth-conditions of the sort envisaged by Van Inwagen, which rely on the sequential order in which we attribute relations, and which we should of course grant. For example, (i) Harry and Sally enter the relation of being a first individual and a second individual such that the first individual loves the second individual if and only if Harry loves Sally and (ii) Sally and Harry enter the relation of being a first individual and a second individual such that the first individual is loved by the second individual if and only if Harry is loved by Sally. Or, more formally, (i) $\lambda xy\ L(Agt(x), Pat(y))(h, s) \leftrightarrow L(Agt(h), Pat(s))$, and (ii) $\lambda xy\ L(Pat(x), Agt(y))(s, h) \leftrightarrow L(Pat(s), Agt(h))$.[11]

5 Conclusion

By distinguishing between the abundant and the sparse realm we can accommodate the pressure to acknowledge converse relations, to which Van Inwagen and the Russell of *PoM* are sensitive, and the opposite pressure to reject them, which is well exemplified in Russell's *ToK*. There is no place for converse relations in the sparse realm, and o-roles are best fit to dispense with them at that level. But we can find room for converse relations in the abundant realm. Once converse relations are admitted, even if only at the abundant level, the problem about naming relations and corresponding converses pointed out by Van Inwagen must be addressed. But by taking seriously the sequential order with which in thinking,

11 What I am proposing here seems to me analogous to what I find in Jan Plate's unpublished manuscript "Basic Positionalism," in that he distinguishes between basic and non-basic relations (I have sparse and abundant relations) and invokes "roles" only for the former. It should be noted, however, that Plate's roles are more like Russell's positions than like my o-roles. I wish to thank Jan for a very useful email exchange on this matter.

writing and speaking we assign o-roles to the relata to which we attribute a relation this problem can be satisfactorily solved.[12]

Bibliography

Castañeda, H. -N. (1975), "Relations and the Identity of Propositions," in *Philosophical Studies*: 28, 237–244.

Castañeda, H. -N. (1982), "Leibniz and Plato's *Phaedo* Theory of Relations and Predication," in *Leibniz*: *Critical and Interpretative Essays*, edited by M. Hooker, Manchester: Manchester University Press, 124–159.

Hochberg, H. (1987), "Russell's Early Analysis of Relational Predication and the Asymmetry of the Predication Relation," in *Philosophia*: 17, 439–459.

MacBride, F. (2016), "Relations," in *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), edited by E. N. Zalta, URL = <https://plato.stanford.edu/archives/win2016/entries/relations/>.

Orilia, F. (2008), "The Problem of Order in Relational States of Affairs: A Leibnizian View," in *Fostering the Ontological Turn. Essays on Gustav Bergmann*, edited by G. Bonino and R. Egidi, Frankfurt: Ontos Verlag, 161–186.

Orilia, F. (2014), "Positions, Ordering Relations and O-Roles," in *Dialectica*: 68, 283–303.

Orilia, F. and Swoyer, C. (2017), "Properties," in *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), edited by E. N. Zalta, URL = <https://plato.stanford.edu/archives/win2017/entries/properties/>.

Russell, B. (1903), *The Principles of Mathematics*, London: George Allen & Unwin.

Russell, B. (1984), *Theory of Knowledge*, edited by E. R. Eames in collaboration with K. Blackwell, London: Routledge.

Van Inwagen, P. (2006), "Names for Relations," in *Philosophical Perspectives*: 20, 453–477.

12 I thankfully acknowledge that my interpretation of Van Inwagen's arguments has been heavily influenced by several conversations and email exchanges with Fraser MacBride. This paper is derived from my presentation at the conference *Quo Vadis, Metapysics?, dedicated to Peter Van Inwagen to commemorate his 75th birthday, Warsaw, September 26th-29th, 2017*. I wish to thank the audience, in particular Peter Simons and Peter Van Inwagen, for their useful comments. Van Inwagen urged me to deal with his 'being to the north of' example, which I had neglected in my presentation.

Andrea C. Bottani

Properties, Nominalisms and Things That Can Be Said

Abstract: According to van Inwagen's theory of the property role, properties are 'things that can be said of things'. Although this theory is, as its proponent says, 'very nearly vacuous', it has in his view an impressive list of substantive consequences about the nature of properties – in particular, it entails that properties are abstract and universal, and thus that nominalism is false. In this paper, I argue that 1) the very idea of a thing that can be said of things is less clear than van Inwagen seems to believe, since it can admit of two different interpretations; 2) in one of these interpretations, which is perfectly coherent and defensible even though it is not van Inwagen's preferred one, the idea fails to entail that properties are universal, and it is far from clear that it entails that properties are abstract. Therefore, the idea that properties are things that can be said of things does not imply Platonism about properties.

1 Nominalisms

In the ongoing philosophical discussion, 'nominalism' is currently used to refer to one or more of the following three theses: A) nothing is universal; B) nothing is abstract; C) nothing is a property. These are logically independent assumptions, since each of them can be coherently endorsed while the other two are rejected, and each pair of them can be coherently endorsed, at least in principle, while the remaining one is rejected; also, they can be coherently jointly endorsed, or jointly rejected, all together. To give just a few examples, class nominalists embrace both A) and C), but reject B); immanent universalists, or some of them, embrace B), but reject both A) and C); mereological nominalists embrace both A) and B), but reject C); trope nominalists embrace A), but reject both B) and C); austere nominalists embrace, and Platonists reject, all of them.[1]

In a number of works (among which Van Inwagen (2004a,b, 2011, 2015, 2017)), Peter van Inwagen argues for a *lato sensu* Platonist theory of properties, according to which all of A), B) and C) are false. Interestingly enough, however, it is only reluctantly that he does so. 'It seems to me', he says '[…] that it is perfectly evident that nominalism is to be preferred to platonism' (Van Inwagen (2004a), p. 15). Pla-

1 For these general approaches see respectively, among others, Quinton (1957), Armstrong (1989), Quine (1950), Devitt (1980), Russell (1912).

https://doi.org/10.1515/9783110664812-017

tonists divide reality into two parts, the concrete part to which we belong and the abstract part, whose inhabitants are completely unlike everything in 'our' part. And it is clear, van Inwagen says, that 'it would be better not to believe in the other part of reality, the other category of things, if we could manage it' (Van Inwagen (2004a), p. 15). But unfortunately, he says, we can't. And the reasons why we can't are, in his view, ultimately two. First, we have no alternative but to assume that properties exist; second, properties can only be abstract and universal. The reason why A) and B) must be rejected is thus partly that C) must be rejected and partly that properties cannot have the same nature as what belongs to 'our part' of reality. There is no systematic theoretical reason why Van Inwagen thinks that the existence of properties is undeniable; it is a simple fact, in his view, that quantification over properties is a pervasive, ineliminable and ontologically committing component of our discourse about reality. But the reason why properties cannot belong to our part of reality is for him a certain theory of what he calls the 'property role', namely, the idea that properties, whatever they may be, certainly play the role of 'things that can be said of things'. This does not seem to say much that is substantive about properties, but it has, according to van Inwagen, a number of crucially important consequences for their nature.

It is on the interface between van Inwagen's theory of the property role and his general theory of properties that I intend to focus here. On the one hand, as I want to argue, the very idea of a thing that can be said of things is less clear than van Inwagen seems to assume. On the other hand, what his theory of the property role seems to entail about the metaphysical nature of properties is less substantive than it is supposed to be. In particular, as I shall argue, van Inwagen's theory of the property role can ground the rejection of neither A) nor B).

2 Property role and material adequacy

What is van Inwagen's theory of the property role? The theory, he says,

> [...] identifies the property role with the role "thing that can be said of something." This role is a special case of the role "thing that can be said." Some things that can be said are things that can be said period, things that can be said full stop. For example: that London has a population of over seven million is something that can be said; another thing that can be said is that no orchid has ever filed an income-tax return. But these things – 'propositions' is the usual name for them – are not things that can be said of anything, not even of London and orchids. One can, however, say of London that it has a population of over seven million, and one can also say this, this very same thing, of New York. And, of course, one can say it of Mexico City and of Oxford. (It can be said only falsely of Oxford, of course, but lies and honest mistakes are possible.) I will assume that anything that can be said of anything can

> be said of anything else. [...] Let us call such things, propositions and things that can be said of things, assertibles. The assertibles that are not propositions, the things that can be said of things, we may call unsaturated assertibles. [...] I propose, therefore, that properties be identified with unsaturated assertibles, with things that can be said of things. [...] what is the property whiteness but something we, in speaking of things, occasionally predicate of some of them? And what is predicating something of something but saying the former of the latter? (Van Inwagen (2004a), pp. 27–28)

As far as I can see, a theory of the 'property role' is not a theory of properties, for it is not a theory of their nature (although it may yield robust information about their nature). It rather looks like a sort of touchstone for the material adequacy of any theories of properties. 'Property' shares with 'truth' an important feature: it is a word used by both philosophers and common people. Tarski's material adequacy condition for truth definitions was intended to give a criterion to decide when a definition is just a definition *of truth* in our sense, and not of something else. Even if van Inwagen does not say so, his theory of the 'property role' (henceforth, TPR) can be interpreted as having a similar status. It can be used to give a general criterion to decide when a philosophical theory is just a theory *of properties* in our sense, and not of something else. One way of stating such a criterion of adequacy might run as follows: either a theory of properties *is consistent* with TPR or it is materially inadequate. Another way might be: either a theory of properties *entails* TPR or it is materially inadequate. I will stick to the former, weaker version of the criterion. Of course, whatever they may be, properties can be predicated by someone of something; so, we cannot pretend that properties are things of a sort such that things of that sort cannot be predicated of anything.

Treating TPR as a sort of material adequacy condition for theories of properties also helps us to understand why TPR is, as its proponent says, 'very nearly vacuous'. Very often, however, *prima facie* neutral material adequacy conditions have *prima facie* hidden substantive consequences. For, if some materially adequate theory incorporates a substantive assumption *s*, and no materially adequate theory can fail to incorporate *s*, then *s* can be said to be a consequence of the material adequacy condition. Tarski's material adequacy condition for truth definitions, for example, seems to entail *prima facie* hidden substantive information about the relation between truth and reference.[2] Likewise, if van Inwagen is right, his 'very nearly vacuous' TPR yields an impressive list of substantive consequences about the nature of properties. Here is a schematic list:

2 See Tarski (1936).

1) Properties are universal (since very often the same thing can truly be said of more than one thing).
2) Properties are abstract (because they stand to one-place open sentences as propositions stand to closed sentences, which no concrete entity can do. So, they are not spatiotemporally located and do not enter into causal chains).
3) Properties cannot be constituents of objects; objects are 'blobs' with no ontological structure (from 2).
4) Particulars are related to their properties by instantiation (exemplification, having); so, particulars and properties can only find place in a polycategorial, relational ontology (from 1 and 3).[3]
5) Existence is a property (for one can say of something that it exists).
6) Properties cannot be perceived (from 2).
7) Haecceities are properties (for one can say of something that it is identical to itself).
8) Properties are abundant, not sparse (for not everything that can be said of things carves nature at its joints).
9) Properties can be uninstantiated and even *necessarily* uninstantiated (for we can say of things inconsistent things).
10) Properties necessarily exist (from 9, and the assumption that accessibility is symmetrical).
11) In no way are properties more basic than the things that have them (from 3).

While TPR does not constitute a theory of properties, this list of theses certainly does. But, as I said, there is a sense in which each of the theses in the list can be said to flow from TPR. This is that the denial of any one of them is purportedly incompatible with TPR, and thus any materially adequate theory of properties must include all of them.

3 Things that can be said. Adverbialism

There are however two problems with the list. One is that it seems too long, for some theses in the list do not seem to flow from TPR in the intended sense. In particular, as I shall argue, it is not clear that 1), and perhaps 2), are indeed consequences of TPR; if they fail to be, TPR is not enough to reject nominalism in the versions A) and B). The other problem with the list is that it also seems too short, for there are other significant theses that seem to flow from TPR in the intended sense no less than 1)–11) supposedly do. Let me consider the latter problem first and the former later.

3 About 2) and 3) see van Van Inwagen (2011).

One might deny that the latter is a problem at all. After all, van Inwagen has never taken the list of 1)–11) as exhaustive; on the contrary, he says that "there are many other interesting and important theses about properties than those I have considered" (Van Inwagen (2011), p. 138). But one problem is that, among the assumptions that have not been put in the list, there is the denial of a thesis that van Inwagen defended a number of years ago. This is the idea that ordinary objects like persons, cats and trees, can have and fail to have the same property at different times without having 'temporal parts'. This thesis is widely known as 'adverbialism', and a number of different versions of adverbialism have indeed been in circulation. Here is van Inwagen's favorite one.

> [...] when we say that Descartes was hungry at t^i, we are saying [...] that it bore the time-indexed relation having-at-t^i to hunger. (Van Inwagen (1990), p. 247).

> One may say both that the relation "x has F at t" is primitive and that its connection with "x has F" is not inexplicable. One need only maintain that "x has F" is the defined or derived relation, and "x has F at t" the undefined or primitive relation. [...] And I do maintain this. To say that Descartes had the property of being human is to say that he had that property at every time at which he existed. To say that he had the property of being a philosopher is to say that he had that property at every member of some important and salient class of moments-his adult life, say. (Van Inwagen (1990), p. 250)

Adverbialism, I believe, is incompatible with the idea that properties are what we say of things. Suppose someone (call her 'speaker A') says that Descartes was thirsty at time t^n, where t^n is (say) 8.00 pm on October 4, 1648. What is it that one says of Descartes if one says that? The obvious answer is: that he was thirsty at time t^n', what else? Thus, if properties are things that can be said of things, the property that speaker A ascribes to Descartes by saying what she says is *being thirsty at time* t^n. But, according to van Inwagen's adverbialism, the property that speaker A ascribes to Descartes is *being thirsty* full stop, and the proposition is true just in case Descartes has this property relative to time t^n. Therefore, adverbialism entails that, when speaker A says what she says, the property she ascribes to Descartes (*being thirsty*) is not what she says of him (*that he was thirsty at* t^n). Suppose someone else – speaker B – says that Descartes was thirsty at time t^k, where t^k is 11 am on August 14, 1646. Of course, speaker A and speaker B do not say of Descartes the same thing. But, according to adverbialism, the two speakers ascribe the same properties to Descartes, that is, *being thirsty* full stop. And the reason why one can tell the truth and the other fail to do so is that Descartes can have the property of being thirsty at t^n and not at t^k (or vice versa).

One might suggest that what they do is to say the same thing (being thirsty) of two different couples (Descartes and a time in the one case, Descartes and an-

other time in the other case). If this were so, however, the ascribed property would be two-place (*being thirsty at*…), which is in conflict with adverbialism. According to the adverbialist, indeed, instantiation is time indexed, while the ascribed property of being thirsty is non-temporalized. So, if adverbialism is true, what the speaker says of Descartes is not the property she ascribes to him.

Of course, this is not a general proof that properties are not what we say of things (especially if one does not embrace adverbialism). It is rather a merely *ad hominem* argument to the effect that, if one believes that properties are what we say of things, then one cannot believe that instantiation is time-indexed (i.e., that things can only instantiate a property relative to a time).

There is however in adverbialism a general intuition of which it is not easy to get rid of. It is that, when one says 'Trump is rich', the copula contributes neither to picking out what one speaks of nor what one says of it, but instead to ascribing the latter thing to the former. Is this intuition compatible with TPR? I shall shortly return to this later.

4 Things that can be said of things. Their individuation

There is another important corollary of the idea that properties are things that can be said of things that is not in the list. Van Inwagen says nothing about the consequences of TPR, if any, for the individuation of properties, i.e., for their identity criteria. But he *could* say something. For, although the idea that properties are things that can be said of things suggests no precise criterion of identity for properties, there is at least one general consequence that the idea seems to have for the adequacy of any such criterion: if properties are things that can be said of things, then properties can only be individuated hyperintentionally. Why? There may be more than one reason. One is that, on van Inwagen's view, (i)–(iii) entail that there are necessarily uninstantiated properties – say, being a square circle and being both 40.000 and 10.000 kms long. But, certainly, what we say of something by ascribing to it the former property is not what we say of it by ascribing to it the latter property. So, necessarily coextensional properties are possibly different.

There is another less visible but very important reason to think that unsaturated assertibles are possibly different even if they are necessarily coextensive. It is that, if TPR is true, then particulars stand in a relation of instantiation (exemplification, having) with their properties. So, imagine I say of Donald Trump that he is rich and then I say of him that he has the property of being rich. How many things would I have said of him? One or two? For those who believe that necessary

coextensionality entails property-identity, the only possible answer is 'one'. For, necessarily, for any *x*, *x* has the property of being rich just in case *x* has the property of having that property; so, these properties are necessarily coextensional, that is, they are had by the same things in every possible world.

But property *P* ought not to be the same as the property of having *P*, otherwise there could be no 'external' relation of having between things we say and things of which we say them, and van Inwagen's theory of properties would be incompatible with a relational ontology (see point 4 above). *Having* (*instantiation*, *exemplification* or whatever word you prefer to use), would be inside what we say of things, and it couldn't connect what we say of things with anything else. If richness consisted in bearing the relation of having to richness – that is, in bearing the relation of having to having richness, and so on and so forth *ad infinitum* – then no relation of having could connect Trump to richness. To say that Trump stands in the relation of instantiation with the property of being rich would be just a roundabout way to say of him that he is rich.

If TPR is true, then, properties must be hyperintensionally individuated. In particular, for any property *P*, *P* and the property of having *P* must be different even though necessarily coextensional. This yields two undesirable consequences. First, for any property, there is an infinite number of necessarily coextensional properties. This has nothing to do with what is known as the 'Bradley regress', for van Inwagen does not think that instantiation of higher level properties such as *having richness* explains in any way instantiation of lower level properties such as *richness*; and it is no violation of what Lewis calls *qualitative* economy, but just of the more innocuous quantitative economy. Despite that, one might still have huge difficulties in grasping what the property of *having having having having richness* may be and what it is that distinguishes it (say) from the property of *having having having richness*.

Second, if things that can be said of things are hyperintentional creatures, it is far from easy to see what keeps them apart from the senses of the expressions we use to say these things. And this, one might suspect, potentially ends up conflating semantics and ontology.

Curiously enough, Van Inwagen uses quotation marks to refer to unsaturated assertibles. 'That it has a population of over two million' and 'that it once filed an income-tax return' are thus for him examples of unsaturated assertibles. This is unusual, for quotations marks are currently employed to refer to expressions (conceived of as either types or tokens), and unsaturated assertibles certainly are neither expression tokens (for they are purportedly abstract) nor expression types (for the same expression type can have different meanings in different languages or discourses). Although the ambiguity could easily be avoided

by replacing quotation marks with other symbols – say, asterisks[4] – a strange fact remains: whatever expressions are used to pick out unsaturated assertibles ('that…', *that…* or whatever), these expressions can never be like 'the Morning Star' and 'the Evening Star' in having the same referent and different Fregean senses. Although this tells a lot about the identity of properties, it does not tell enough to give a working identity criterion for them. Here is why.

Suppose there is something one says of things when one says of them that they are α and there is something one says of things when one says of them that they are β. Are there any necessary and sufficient conditions for the former thing – the former property – be identical with the latter? If they were identical just in case, necessarily, for any x, x is α just in case x is β, properties would be identical just in case they are necessarily equiextensional. In such a case, however, they would fail to be individuated hyperintensionally. If properties are hyperintensional creatures, their criterion of identity can only be grounded on synonymy: the thing one says of things when one says of them that they are α and the thing one says of things when one says of them that they are β – these properties – are the same property just in case 'α' and 'β' have the same Fregean sense. Unfortunately, this is not a good criterion of identity for properties, because it involves a confusion between use and mention. Can the confusion be removed from the identity criterion and, if so, how? Here is a possible way: for any predicates α and β, any name 'n' the thing one says of the referent of 'n' when one says 'n is α' and the thing one says of it when one says 'n is α' are the same thing just in case 'α' and 'β' have the same Fregean sense. This formulation removes every confusion between use and mention from the criterion, but only at the cost of transforming a criterion of identity for properties (however badly stated) into something else. The reason is that, as van Inwagen remarks, there are innumerable properties for which nobody has ever had a predicate. To any such properties – things that might be said of things but have never been said of anything, for there is no predicate for them – the above criterion of identity can only fail to apply. There is nothing surprising in this: after all, this is a criterion of coreferentiality for expressions denoting properties, not a criterion of identity for properties (compare '*for any sons* s^1 *and* s^2, *the women of whom* s^1 *and* s^2 *are sons are the same woman iff* s^1 *and* s^2 *have a sufficiently similar DNA*', which is a criterion of identity for women that have sons, but not for women in general). No precise criterion of identity for properties can thus follow from TPR, the thesis that properties are things that can be said of things. Nevertheless, as I have said, TPR ties what we say of things

4 **P** would thus refer to the sense of the expression that '*P*' refers to.

(namely, properties) so closely to the senses of the expressions we use to do that, as to apparently conflate semantics and ontology.

Can TPR be retained without conflating the two levels? It depends, I believe, on TPR's intended meaning. TPR can have more than one intended meaning because the notion of a thing that can be said of things can have more than one intended meaning. Let me explain why. Suppose one names something, for example Donald Trump, and then says some words about him, for example 'is rich'. What does one say of Trump if one says of him 'is rich'? The question is ambiguous, for it can be taken as being either about the sense or about the reference of 'is rich' (either about what is ascribed to Trump by means of 'is rich' or about the way 'is rich' makes the ascription). TPR, the thesis that properties are things we say of things, is ambiguous in the same way. Either the things we say of things are taken to be the senses of the words we use when we say them of things; or they are taken to be the entities we say of, predicate of, or ascribe to things as a consequence of saying these words. In both cases, properties are things we can say of things, but in different senses.

One might thus say: TPR is true, for properties are indeed things that can be said of things (they can be said of things for they can be *ascribed* to things). But they are not the meanings of our words, even though they can only be ascribed to things by using meaningful words.[5] Likewise, territories can only be ascribed to nations – and honors can only be ascribed to persons – by using meaningful words, but territories and honors are not meanings (even less are they words, either taken as expression types or as expression tokens). Properties are said of things insofar as they are ascribed or predicated of things by means of words, not insofar as they are the senses of the words by means of which they are said of things. This reinterpretation of TPR does not exclude that properties are 'unsaturated'. To see why, consider Frege's theory of *Sinn* und *Bedeutung*. On Frege's account, predicates have both a *sense* and an unsaturated referent (a 'concept' conceived of as a function from objects to truth values), and predicates denoting the same concept can nonetheless have different senses (which is, for example, the case of 'having four sides' and 'having four angles'). Neither does this interpretation of TPR exclude that properties are 'assertibles', for there are different senses in which words denoting sets, tropes, immanent universals and mereological wholes can all be said to be true of particular objects.

5 This may perhaps reconcile TRP with the adverbialist intuition that the copula contributes neither to picking out what one speaks of nor what one says of it, but instead to ascribing the latter thing – the property – to the former.

In this intended meaning, TPR is perfectly compatible with a sharp distinction between the ontological and the semantic, for properties are no longer hyperintentionally individuated from this point of view, and their identity no longer consists in the synonymy of the expressions used to ascribe them to objects. But this is not van Inwagen's interpretation of TPR. And one might doubt that TPR continues to yield 1)–11) as consequences once it is reinterpreted in this way.

5 Things that can be said of things. Mereological nominalism

I've explained why I believe that van Inwagen's list of corollaries of his 'nearly vacuous' theory of property role might be longer. Let me say now why I believe that it should perhaps also be shorter – i.e., why I think that some purported consequences may fail to follow from the nearly vacuous theory, especially once it is reinterpreted as I have indicated. I want to concede for the sake of argument that, if properties are things that can be said of things, then they cannot be constituents of objects.[6] But I deny that, if properties are things of that sort, they can only be universal; and I doubt that, if they are things of that sort, they can only be abstract. My reason to deny the former assumption and doubt about the latter is that I think there can *in principle* be theories that deny that properties are universal, but that are compatible with the idea that properties are things that can be said of things and are not constituents of objects. And I do not see conclusive reasons to exclude that there can be theories that are compatible with both these assumptions but that deny that properties are abstract.

One such theory is perhaps mereological nominalism, according to which the property of redness is the big red object, namely the total sum of red objects, and being red is being a part of that object. The big red object is neither universal nor abstract; and certainly it is not a component of any London red phone box, or of any red fire extinguisher. Is there any sense in which it can be said of a red phone box or a red fire extinguisher? Maybe so. According to both van Inwagen and the mereological nominalist, 'this fire extinguisher is red' is true just in case 'is red' is true of this fire extinguisher. According to van Inwagen, however, this is the

6 See Van Inwagen (2004b), pp. 134–135: "If this pen exists, there are no doubt lots of things that are in some sense its parts or constituents: atoms, small manufactured items… […] But "that it is a writing instrument", although it can be said truly of the pen – and is thus, in my view, one of the properties of the pen – is not one of the parts of the pen."

case iff the fire extinguisher instantiates a universal abstract entity called 'that is red', while, according to the mereological nominalist, this is the case iff the fire extinguisher is part of the red big object.

Mereological nominalism runs into well-known difficulties in dealing with countable properties such as *being a horse*, or *being a chair*. Leaving that aside, there is at least one case in which the property of being red is a component of a red thing according to mereological nominalists: the case of the big red object itself, which is of course an improper part of itself. Thus, even if the big red object could be said of things in some sense of 'say of things', mereological nominalism is definitely not a good example of a theory according to which properties are never components of objects – and thus, *a fortiori*, not a good example of a theory according to which properties are particular and concrete, can be said of objects and can never be components of things.

6 Things that can be said of things. Locationism about properties

Perhaps, a good example of such a theory is locationism, the idea that properties are *sui generis* locations, i.e., places of a metric quality space, and having properties is being located in such places (having mass, for example, is being located somewhere on a line; being exactly 2 grams is staying on a precise point of the line; and being at least 2 grams is staying somewhere on a segment of the line).

Although it has been seldom if ever fully developed, the idea of a quality space is not entirely new and can already be found in Van Fraassen (1967) and Stalnaker (1979), who reworks van Fraassen's proposal. Here is a passage from Stalnaker:

> Properties, on van Fraassen's account, are represented by regions of a logical space, or quality space. For example, the color spectrum might be a dimension of such a space; the color red would then be identified with a region defined by a segment of the color dimension. The temperature scale might be another dimension. The relation warmer than would be identified with a set of ordered pairs of points of the space defined by this dimension… The language [of the model] is interpreted by assigning to each one-place predicate a subset of the points in logical space. Such a predicate is satisfied by an individual just in case the location function locates the individual at one of the points included in the value assigned to the predicate. If the predicate is to mean is red, then its value will be the region of logical space defined by the relevant segment of the color dimension. The open sentence "x is red" will be satisfied by an individual just in case the individual is located somewhere in that region – just in case the location function colors it some shade of red. This account generalizes to cover n-place predicates and n-ary relations in the obvious way. (Stalnaker (1979))

There may be different ways to establish the details of the account, but this is the general idea. Van Fraassen and Stalnaker, however, seem to understand quality space more as a visual representation of properties than as a theory of what they really are. Hawthorne and Sider (2002) also defend a version of the idea, but define places as n-tuples of particulars, which seems to deprive the idea of any genuinely locationist content. Gärdenfors (2004) gives a systematic account of the structure of quality space, but is fully explicit about treating it as a mere conceptual construct, which seems to make his view a version of a concept theory of properties. Likewise, Funkhouser (2006) uses spatial models to represent determinables such as *mass*, determinates such as *1 kilo mass* and their relations. As far as I know, Cowling (2014) is the first to take quality space as a fundamental ontological structure. Cowling examines in depth the consequences of the idea for instantiation, but is neutral between relationalism and absolutism about quality space, and so he does not say much about the ultimate nature and the metric structure of quality space.

I do not want to give anything more than a rough idea of locationism about properties, but only to explain why I think that locationism is perhaps a coherent example of a theory of properties which takes properties to be things that can be said of things, but nonetheless are neither universal nor abstract nor components of objects. As to locationism, let me limit myself to the following schematic clarifications. Locationism conceives of a determinable quantity such as mass as a line, determinate masses as points (sometimes occupied, sometimes not) on the line and mass distances as line segments having mass points as limits. Having a determinate mass is being located at a particular point on the line. Having different masses is being at the extremities of a segment of the line. Having *some mass* is being located somewhere on the line. Unoccupied points are uninstantiated determinate masses. But mass is just one dimension of an n-dimensional quality space: there is mass, charge, size, speed, volume, density, temperature etc. Each of them is a line, each line is a dimension of a unitary quality space. Every natural property is an extended or punctual region in quality space. Having a property is having a location in quality space. Locationism takes quality space at face value, as something that enjoys objective existence and not as a model or representation of what properties are. The super-determinable property of having mass, for example, is assumed *to be* that line, and the various determinate masses are assumed to be points on that line. Therefore, having a mass *just is* being located in a place. Just as one can defend relationalist or absolutist theories of physical space, one can embrace relationalist or absolutist theories of quality space. You can assume that the manifold of quality points and regions exist independently from the particulars that are located in it, or instead believe that it is grounded in the totality of the existing particulars (somewhat as the manifold of spacetime points

and regions are grounded in the totality of the existing particulars in Leibniz's *Analysis Situs*[7]).

Now, consider this general image and ask: are properties conceived that way – that is, as quality places – universal? Are they abstract? Are they components of the particulars that are located in them? And, last but not least, can they be said of things?

To begin with, physical regions of spacetime are not components of things located in them (it would be a category mistake to believe otherwise. And there is no reason to think that quality regions compose the things located in them anymore than spacetime regions do. So locationism is inherently relational (instantiation is location), and is incompatible with any constituent ontology.

Second, regions of quality space can be neither Platonic nor immanent universals. They are not Platonic universals because they do not transcend the physical world any more than regions of spacetime do (physical objects are located in an n-dimensional box, where some dimensions are spatiotemporal and some qualitative). And they are not immanent universals, because they cannot be found in spacetime. Wondering whether they do, looks like wondering whether latitudes are in longitudes, or whether places are in time, or whether the horizontal axis is in the vertical axis. In all these cases, some dimension(s) cross(es) some other dimension(s), but it would be completely wrong (perhaps meaningless, just a category mistake) to say that either dimension is inside the other. Although a quality region is not universal, a number of distinct particulars can be located in it, just as a number of distinct particulars can have the very same universal property. Although a qualitative region is non-universal, it can be measured in many ways, just as a spatial region can.

Are regions of quality space abstract? In some sense of 'abstract', they seem indeed to be abstract. First, they are not located in spacetime, as I said earlier. Second, they cannot enter in causal chains at all, or so it seems. Third, they cannot be perceived. So they seem to be as distant from concrete particulars as anything could be. But we do not think of regions in physical spacetime as abstract entities, yet they seem to behave much more like quality space regions than like concrete particulars. So, in what sense, if any, are quality regions more abstract than spatiotemporal ones? Just like quality space regions, they do not seem to enter in causal chains at all (perhaps things located in them can enter into causal chains in virtue of being located in them, but spatiotemporal regions themselves cannot (which also seems to be the case with quality space regions). And, just

7 One important difference is that Leibniz was anti-realist or phenomenalist about physical space (see Leibniz (1978), II).

like quality space regions, spacetime regions cannot be perceived (one can perceive things located in regions, not regions themselves). Finally, one might doubt that spatiotemporal regions, unlike quality regions, are located in spacetime, for they do not seem to bear to themselves the very same kind or relationship that concrete particulars bear to them. Certainly, they are parts of other spacetime regions; thus, if spacetime is the total sum of all space regions, they are parts of spacetime. But are they, inasmuch as they are parts of it, also *located* in it? Is any spatiotemporal region exactly located in itself? Insofar as the answer is unclear, it is unclear whether quality space regions are more abstract than their spacetime correlates. But there are perhaps other reasons to believe so.

One reason may be that concrete particulars seem to be more apt to be exactly co-located in quality space than in spacetime. According to some philosophers, material objects are impenetrable in spacetime but possibly penetrable in quality space (there might be objects different *solo numero*, or exactly co-located in quality space, yet distinct like Black's (Black (1952)) iron spheres; but see Hacking (1975) for a reinterpretation of this imagined situation as one in which there is just one sphere in a non-Euclidean space). Perhaps, spatiotemporal distance suffices to allow difference *solo numero*, while quality distance might be insufficient to license spatiotemporal co-location. This is a matter for weighty debate, however, and little if anything is perfectly clear here. On the one hand, Locke and Wiggins argued that ordinary objects of different sorts can indeed occupy the same place at the same time[8], even though imagined situations in which they do so at every moment of their lifespans are barely intelligible. On the other hand, Leibniz famously denied that things can be different *solo numero*, and the discussion is still lively and ongoing.

Another possible reason to treat quality regions as abstract concerns modality. According to David Lewis, the only glue able to keep the many pieces of a single world together is spatiotemporal connection, while properties are sets of elements scattered across the manifold of possible worlds. Hence, the boundaries of a world are the limits of its spacetime, while quality space transcends any single world. Since no sum of different possible worlds can be concrete inasmuch as it is spatiotemporally disconnected, the concrete is the spatiotemporal, the world-transcendent is the abstract. It is not clear, however, why a sum of concrete possible worlds should be abstract. Suppose what we call real is strangely scattered in a number of spatiotemporally disconnected fragments. Why should such a sum

8 See Locke (1975), Chapter XXVII of Book II, Wiggins (1968). Hughes (1997) argues that, in some cases, two different things can be in the same place at the same time even if they are of the very same sort.

deserve to be called 'abstract' despite the fact that all its parts are concrete?[9] Inasmuch as this is far from clear, no decisive reason remains to take it quality regions to be more abstract than spacetime regions.

The important question is: are quality space regions things that can be said of things? Can they be treated as properties in the sense of TPR? Certainly, quality space regions cannot be taken to be the Fregean senses of the words we use when we say things of things. But there is nothing strange in taking them to be what we ascribe to things by saying those words. Sure, one may well feel uncomfortable at hearing that, when one says that a rod has a mass of 2 kilos, one is ascribing to the rod some coordinate on an invisible 'line of mass' – namely, a certain quality location. But imagine that one has never heard of Platonic properties, and universals, and so on: would one not feel the same sense of bewilderment at hearing that, when one says that a rod has a mass of 2 kilos, one is ascribing to the rod an immaterial transcendent, invisible universal that the rod instantiates in some sense of this word? Maybe much of the bewilderment disappears if you are allowed to call the universal immaterial entity "*having a mass of 2 kilos*". What is far from clear is why the locationist could not give the same name to a specific region of her quality space.

According to locationism, then, properties 1) are not components of things, 2) they are not universal, 3) there is no conclusive reason to think that they are abstract and 4) they are things that can be said of things (even if they cannot be taken to be the senses of the words we use when we say things of things). If this is true, a theory of properties according to which properties are not components of objects can be materially adequate – which is to say, compatible with TPR – even though it denies that properties are universal and (perhaps) that they are abstract. Therefore, the conjunction of TPR and the denial that properties are constituents of things does not entail that properties are universal, and there is no conclusive evidence that it entails that properties are abstract.

It would be better, van Inwagen says, to be nominalists, if only it were possible. I have tried to argue that neither the conception of properties as unsaturated assertibles nor the adoption of a relational ontology, nor even their conjunction can definitely rule out nominalism for those who prefer not to be Platonists.

9 See Van Inwagen (2004b), p. 129.

Bibliography

Armstrong, D. M. (1989), *Universals. An Opinionated Introduction*, Boulder CO: Westview Press.
Black, M. (1952), "The Identity of Indiscernibles", in *Mind*: 61, 153–164.
Cowling, S. (2014), "Instantiation as Location", in *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*: 167, 667–682.
Devitt, M. (1980), "'Ostrich Nominalism' or 'Mirage Realism'?", in *Pacific Philosophical Quarterly*: 61, 433–439.
Frege, G. (1892), "Über Sinn und Bedeutung", in *Zeitschrift für Philosophie und philosophische Kritik*: 100, 25–50.
Funkhouser, E. (2006), "The Determinable-Determinate Relation", in *Nous*: 40, 548–569.
Gärdenfors, P. (2004), *Conceptual Spaces*, Cambridge MA: MIT Press.
Hacking, I. (1975), "The Identity of Indiscernibles", in *Journal of Philosophy*: 72, 249–256.
Hawthorne, J. and Sider, T. (2002), "Locations", in *Philosophical Topics*: 30, 53–76.
Hughes, Ch. (1997), "Same-Kind Coincidence and the Ship of Theseus", in *Mind*: 106, 53–67.
Leibniz, G. W. (1978), *Die philosophischen Schriften von G. W. Leibniz*, edited by C. I. Gerhardt, Hildesheim: Georg Olms.
Lewis, D. (1986), *On the Plurality of Worlds*, Oxford: Blackwell.
Locke, J. (1975), *An Essay concerning Human Understanding*, edited by P. H. Nidditch, Oxford: Oxford University Press.
Quine, W. V. O. (1950), "Identity, Ostension and Hypostasis", in *The Journal of Philosophy*: 47, 621–633.
Quinton, A. (1957), "Properties and Classes", in *Proceedings of the Aristotelian Society*: 50, 33–583.
Russell, B. (1912), *The Problems of Philosophy*, Oxford: Clarendon Press.
Stalnaker, R. (1979), "Anti-Essentialism", in *Midwest Studies in Philosophy*: 4, 343–355.
Tarski, A. (1936), "Der Wahrheitsbegriff in den Formalisierten Sprachen", in *Studia Philosophica*: 1, 261–405.
Van Fraassen, B. (1967), "Meaning Relations Among Predicates", in *Nous*: 1, 160–179.
Van Inwagen, P. (1986), "Two Concepts of Possible Worlds", in *Midwest Studies in Philosophy*: 11, 185–213.
Van Inwagen, P. (1990), "Four-Dimensional Objects", in *Nous*: 24, 245–255.
Van Inwagen, P. (2004a), "Properties", in *Knowledge and Reality. Essays in Honor of Alvin Plantinga*, edited by T. Crisp, M. Davidson and D. Vander Laan, Doordrecht: Springer, 15–34.
Van Inwagen, P. (2004b), "A Theory of Properties", in *Oxford Studies in Metaphysics*: 1, 107–138.
Van Inwagen, P. (2011), "Relational vs. Constituent Ontologies", in *Philosophical Perspectives*: 25, 389–405.
Van Inwagen, P. (2015), "Against Ontological Structure", in *The Problem of Universals in Contemporary Thought*, edited by G. Galluzzo and M. Loux, Cambridge: Cambridge University Press, 46–64.
Van Inwagen, P. (2017), "Concluding Meditation", in *Being, Freedom, and Method*, edited by J. Keller, Oxford: Oxford University Press, 343–394.

Wiggins, D. (1968), "On Being in the Same Place at the same Time (with one Remark about Categories and Materialism)", in *Philosophical Review*: 77, 90–95.
Williams, D. C. (1953), "On Being in the Same Place at the same Time (with one Remark about Categories and Materialism)", in *On the Elements of Being: I*: 7, 3–18.

Christian Kanzian

Paraphrase: A (More or Less) Van Inwagenian Way toward (Moderate) Nominalism

Abstract: The aim of this article is to obey Peter Van Inwagen's imperative – see *Existence* (2014), p. 153: We should wish not to be Platonists if it's rationally possible – and hence work on the if-clause in order to make clear that it is actually rationally possible to avoid Platonism, at least to a significant extent.

What I am interested in is the idea of translating or explaining away of sentences, by means of which we commit ourselves (or seemingly commit ourselves) to abstract objects, for which we have a technical term: paraphrase. What exactly is a paraphrase? How can we distinguish between successful and unsuccessful paraphrase?

The possibility of giving a paraphrase cannot be decided absolutely or simpliciter, but only relative to a toolbox of ontological prerequisites. I intend to present such a toolbox, in order to paraphrase abstract object-sentences, particularly natural kind sentences.

1 Introduction

"It would be better not to believe in abstract objects if we could get away with it." This is the heading of the first section of chapter 8, "A Theory of Properties", of Peter van Inwagen's book *Existence* (Van Inwagen (2014), p. 153). Every philosopher should meditate often on this heading: on the wise advice it includes and the legitimate concern it expresses – a concern which, a few lines below, is also formulated as an imperative: "… a philosopher should wish not to be a Platonist if it's rationally possible … not to be a Platonist." The aim of this article is to obey this imperative: We should wish not to be Platonists if it's rationally possible, and hence work on the if-clause in order to make clear that it is actually rationally possible to avoid Platonism, at least to a significant extent.

We know that Peter van Inwagen is not so optimistic about this – but nevertheless he wants us to try. This is the first reason why the present attempt to promote nominalism can be called "more or less" van Inwagenian. Another reason, which stresses the "more", is that this project operates at the borderline between ontological and metaontological reasoning. Many fights against abstract objects have already been lost in the battlefields of metaontology. For example: two recent metaontological positions, Meinongianism and "Easy Ontology", would find it hard to return to nominalism than would a standard metaontology in the tradi-

https://doi.org/10.1515/9783110664812-018

tion of Quine. This justifies the hope of finding a way toward a nominalism of a van Inwagenian style, because – as is well known – Peter van Inwagen has nothing to do with the above mentioned metaontologies, but is actually the leading Quinean.

According to Quine's metaontology, to exist or to be means to be the value of a bound variable. It is with bound variables that we commit ourselves ontologically. According to van Inwagen, bound variables are nothing more than abbreviations of words and phrases in English uses of such phrases as "it is true of at least one thing …" or, in short, "there are things …". From this we may conclude that our ontology consists of exactly those entities that we mention after "there are …"-expressions, where these expressions are true if the relevant things exist (Cf. Berto and Plebani (2015), p. 20; Van Inwagen (2014), p. 125). The problem is that in many, or perhaps most of our "there are … "-sentences, we do not really intend to commit ourselves to entities, i.e. to basic elements of the world. Some philosophers argue that we sometimes utter "there are …"-sentences outside the *ontological room* (cf. Van Inwagen (2014), p. vii). We utter not only sentences such as "There are sheep on the grass" or (if we favor not only *Existence*, but also van Inwagen's *Material Beings*) "There are some atoms arranged chair-wise in this room", but also sentences such as "There are troubles arising", "There are cloud-ribbons above the mountains around Innsbruck". The decisive question is: with which of these statements do we really commit ourselves ontologically? Are we ready to accept socio-psychological or meteorological phenomena as entities in our ontology in the same way that we accept substances? A provisional answer might be: no. We commit ourselves only with those "there are …"-statements in which the expressions after "there are" cannot be translated away. In the case of the latter statements, there are surely alternative terms to which we can transfer the ontological weight of the original sentences, whereas in the case of the former sentences – at least if we follow van Inwagen – there are not.

My aim is not to spend time on these examples. What I am interested in instead is the idea of translating or explaining away, for which we have a technical term: *paraphrase*. What exactly is a paraphrase? How can we distinguish between successful and unsuccessful paraphrase? And especially, is there a successful paraphrase for sentences, by means of which we commit ourselves (or seemingly commit ourselves) to abstract objects? The if-clause in our "It would be better not to believe in abstract objects if we could get away with it" can also be understood as: "if sentences employing abstract-object terms after 'there are' can be paraphrased into sentences which do not." This paper aims to pinpoint such a paraphrase. But let us proceed step by step, and begin with the principal question of what a paraphrase is.

2 Paraphrase

What exactly is a paraphrase? If we intend to follow van Inwagen, we do well to start where he does, and this is with Quine. In his *Quine's 1946 Lecture on Nominalism* (Van Inwagen (2014), chapter 6), van Inwagen refers to a passage in which Quine introduces the concept "paraphrase" as referring to a method to reformulate a sentence to show that the reference e.g. to abstract objects (such as biological species) is an avoidable manner of speaking (*ibi*d., 127). In the canonical grammar of quantification: We evade the unwanted commitment by showing that the seeming quantification over abstract objects, e.g. biological species, is avoidable (*ibid*., 127f). Take as a starting point our everyday talk about natural kinds, such as talk of spiders. I understand Quine as saying that, even if we use natural-kind terms in subject position (e.g., "Spiders are frightening") outside the "ontological room", it is avoidable to conclude "There is something which is spiderhood and which is frightening" in an ontologically strict sense. His reason, I take it, is that we can reformulate the latter sentence ("there is … spiderhood …") as a sentence in which the "there is" is not followed by any noun denoting abstract objects (e.g. by "there are … particular spiders…").

To clear up a misunderstanding that immediately suggests itself, we can again make reference to *Existence* – in particular to van Inwagen's critical comment on Alston (*ibid*., chapter 7). Alston has formulated a well known dilemma about paraphrases: According to Alston a paraphrase is either, first horn, an adequate translation of the original sentence, or (second horn) it is not. On the first horn, the paraphrase just says the same thing as the original sentence, perhaps in different words. In this case, any ontological commitments made by the original sentence cannot be avoided. On the second horn, in which the translation is not adequate, even if we avoid unwelcome "there is/are…"-sentences, it fails to be a paraphrase. Moreover, the paraphrase cannot be used *to reduce* the entities to which we commit ourselves in the original sentence to entities about which we speak in the translation (*ibid*., p. 138). Van Inwagen rejects Alston by pointing out that his dilemma rests on two questionable premises. The first concerns Alston's understanding of ontological reduction. An ontological reduction can never amount to removing the acceptance from one sort of entity and giving it to another sort – yet Alston seems to suppose that it can. Consequently, a paraphrase does not move us from commitment to one kind of entity to a commitment to another kind. A paraphrase is a vehicle for avoiding *seeming* commitments, not real ones. If it succeeds, then we have no reason to accept the putative kind of entity in the original phrase. This brings us to Alston's second premise, which is that the para-

phrase has the same meaning as the original sentence: the considerations just given show that this premise is false.

In short: to escape Alston's dilemma, it is necessary to understand that the ontological commitment of the original sentence is not a real one (Van Inwagen (2014), pp. 149ff). If we paraphrase a "there is …"-sentence about (say) biological species into a sentence in which the "there is…" is followed by an expression referring to particulars, we do so because we assume that the original sentence, which differs in its meaning from the paraphrase, includes a merely apparent ontological commitment.

With this remark we may take a further step in our attempt to characterize paraphrase. We must answer the question: How can this sort of removal of an ontological commitment succeed? How do we distinguish a successful from an unsuccessful paraphrase? Are there criteria for the success of a paraphrase?

Following Peter van Inwagen, we are in the fortuitous situation of being able to identify a group of promising criteria. The first can be deduced immediately from the general characterization of paraphrase given above. A paraphrase's removal of a merely apparent ontological commitment is successful if "the original sentence seems to imply the existence of so-and-so's, … and the paraphrase does not imply the existence of so-and-so's" (*ibid*., p. 148). To this initial criterion van Inwagen adds another, more common-sense criterion: A paraphrase is successful if "… [the paraphrase] could … be used for all the same purposes as the original in the business of everyday life" (*ibid*., see also *ibid*., p. 161). We can understand this criterion in the following way: Take for example an arachnophobic person who in his everyday life frequently says: "I am afraid of spiders". Some ontologists would interpret this statement as a "there is a natural kind"-sentence along the lines of: "There is something which is a natural kind, spiderhood, which is feared by the speaker". A successful paraphrase of such a sentence could be a statement in which the speaker commits himself not to kinds, but to particular organisms, e.g. "There are some things, which are spiders, which I fear". The crucial point is that the "there are … spiders"[1]-sentence is no less foreign to arachnophobic everday usage than the "there is … spiderhood"-sentence. The paraphrase must have the same "implications for everyday action …" (Van Inwagen (2014), p. 147). In everyday life it could be used for all the same purposes as the original, as well as in specific practices such as biology, psychiatry, and so forth.[2]

1 In what follows I continue using "there are … [e.g.] spiders" as an abbreviation for "there are some things, which are [e.g.] spiders".

2 For a comparable criteriology see von Solodkoff (2014).

A further point can be traced back to Quine, who distinguishes between the content of an ontological theory, and what he calls its *ideology*: undefined primitives which must be presupposed in order to arrive at the theory's intended content. It is important to see that whether one accepts or rejects a paraphrase depends on the ideological component of the ontological theory. We need not go into detail on Quine's concept of ideology or the notion of primitiveness. What matters is that this distinction enables us to derive a further criterion for the success of a paraphrase: A paraphrase is successful if it is implied by reasonable basic ontological presuppositions. These considerations should yield an acceptable interpretation of van Inwagen's abovementioned concern "Let's avoid Platonism, if this is rationally possible": "Let's look for an adequate ideological toolbox in order to paraphrase seeming 'there are abstract objects'-sentences".

Of course, we could start long metatheoretical debates about the notion of adequacy at issue here. *Consistence* and *economy* seem to be uncontroversial components of it. A successful paraphrase must never require accepting mutually excluding primitives or an endless list of undefined concepts. And, if we adhere faithfully to van Inwagen's common-sense criterion, a paraphrase should not be implied by primitives whose acceptance is at odds with everyday action and the relevant scientific practice.

In any case: When a sentence seems to commit us to a kind of entity, a paraphrase is a method of translating that sentence into another sentence that does not commit us to this kind of entity. It is successful if the paraphrase does not imply the existence of the so-and-so's that the original sentence seemed to imply. The theoretical basis of a paraphrase must be consistent, economical, and in accordance with the "practice"- criterion.

3 Paraphrase and the avoidance of abstract objects

3.1 Abstract Objects

With these initial results concerning the concept "paraphrase" and the criteria for the success of a paraphrase in general, we can proceed to the next step: the question of how to paraphrase sentences that commit us (or seem to commit us) to abstract objects. Given the above considerations, our emphasis should be put on the "ideology aspect". Which ontological tools do we need for this sort of paraphrase? How can we apply them to a procedure which also meets the other criteria?

Before we begin, we need a few words about the objects of our reduction via paraphrase: abstract objects. This is not the place to give a proper definition. For present purposes it suffices to introduce abstract objects in contrast with concrete objects (cf. Van Inwagen (2014), pp. 154–158). While the latter have *determinate spatial and temporal characteristics*, the former, abstract objects, lack them. They may lack such characteristics because they exist (if they do) completely outside of our spacetime-system, e.g. in the mode of Platonic or Pythagorean entities; or because their instances occur in space and time, but in and of themselves lack a fixed spacetime position. Some authors allow for this sort of indeterminacy for universal abstract objects (such as colors or natural kinds) that may occur in our space-time-system but in and of themselves have no fixed position in this system. What matters for present purposes is that, if abstract objects exist, they are entities without determinate temporal and spatial positions.

These remarks touch on the problem of classifying abstract objects. There is a possible distinction between universal abstract objects and abstract individuals that I cannot discuss here (cf. Lowe (2006), p. 83), nor can I discuss various other approaches to taxonomizing abstracta. Without any claim to completeness, the paradigmatic cases of abstract objects that I have in mind include mathematical objects such as numbers and sets, natural kinds such as biological species, and universal properties. These are particularly interesting in the present context, because they are actually the preferred objects for attempts of a reduction by paraphrase.

And here is where the rubber hits the road: in attempt to obey van Inwagen's imperative and become moderate nominalists, what we need is to paraphrase sentences in which we (seemingly) speak about the candidates for abstract-entityhood which we have been calling *natural kinds*. reason for this limitation is not only that we should avoid packing too much into a single article, but also the conviction that any attempt to reduce abstract properties and mathematical objects via a paraphrase of abstract properties- and mathematical objects-sentences is ultimately idle. There seems to be no prospect of success for paraphrasing mathematical objects – at least not in the context of the Quine/van Inwagen-metaontology at issue here. When it comes to natural kinds, by contrast, we can immediately deduce a paraphrase of an abstract properties-sentence from a successful paraphrase of a sentence that is seemingly about abstract natural kinds (see section 4.). So, what we need to do is to find a successful paraphrase for "there are …"-sentences containing abstract natural kind terms. In keeping with our discussion above, let us focus especially on the ontological prerequisites needed for such a paraphrase.

3.2 Prerequisites for a paraphrase of natural kinds-sentences

The first concept in our list of prerequisites, *complex unity*, seems at first glance to have nothing to do with natural kinds. The concept "unity" seems intuitively clear, primarily in contrast to "diversity". We can of course differentiate types of unities. On the one hand there are unities that lack any parts whatsoever (in any sense) and can thus be considered *simple unities* (think of Leibniz's monads, for example). On the other hand there are *complex unities*; these are in some way structured, and they have parts (in a liberal sense) or aspects that configure the unity in question. Among such complex unities we can further distinguish between those whose parts are prior, that is, those which consist purely of a number of elements and hence are simple sums of their elements (like heaps of sand), and those whose complex unity is prior to their parts. Unities of this latter sort cannot be derived or reconstructed solely from their elements: They are not simple sums of their parts (in the liberal sense at issue). The notion of this sort of complex unity is the first tool for the intended paraphrase.

The relevance to natural kinds will become clearer when we consider substances that are the paradigmatic members of the (alleged) kinds. According to the most important and influential theory of substances (the Aristotelean one), we can take such substances to be instances of complex unities. Aristotelean substances are unities, but not simple ones. They are instead complex unities: *hyle* and *morphe*, matter and form, are the elements or aspects of their complex structure. The individual matter and the individual form are not parts from which the unity of the whole substance can be derived: Substances are no mere sums of matter and form. In ontological terms, the distinction between matter and form refers to the primary unit of the whole substance.[3] Nevertheless, both the individual matter and the individual form of a substance have irreducible ontological functions: Matter is that *from which* a thing is made. Traditionally, matter is the principle of potentiality. Form, by contrast, is the principle of *how* matter is structured within the complex unity; form is the principle of realization, of *actus*. That is, form is the principle of a substance's *identity* and *individuation*.

The details need not concern us here. What we need for our reduction of substance-kinds are the ideas, first, that substances are complex unities, and, second, that an individual form, which is an irreducible elements of this com-

3 Aristotle explicitly asserts in his *Metaphysics* that it would be a serious mistake to introduce a third element, such as the sum of matter and form, to our notion of the entire substance. Cf. Aristotle, *Metaphysik* Z, 1041b 11–14.

plexity, is – to put it cautiously – relevant for the natural kind of the respective substance.

The second tool in our ideological toolbox is the concept of *ways* (*in which*) *complex unities are*. This tool is borrowed from Jonathan Lowe as well as from John Heil; for the moment we may for the sake of simplicity adopt Lowe's account. He introduces the ways complex unities – or substances – are by appeal to the formal relation of *characterization*. An ontological axiom says that complex unities are something which is characterized: there are ways they are. Characterization, as Lowe understands it, excludes the possibility that substances are *constituted* by the ways they are – that is, he rejects the notion of substance adopted by van Inwagen in "Constitutional Ontologies" (cf. Van Inwagen (2014), chapter 10). Substances are not sums or bundles of the ways they are. Another important point is that ways substances are depend strictly on substances. This particular blueness would not exist if it were not a way this particular book is.

We need not delve any further into a discussion of this dependence, as it would lead us too far afield of the present topic. But we do need discussion of another aspect of substances, which we can call "*derivaty*". The notion of derivaty features most prominently in the work of Lynne Rudder-Baker (cf. Rudder-Baker (2007), pp. 37ff). In the context of the abovementioned complex-unity-account of substances, the claim that substances have "derivaty" means that the characterization of an entire substance derives from the characterization of one of its structural elements, namely individual form or individual matter. In other words, the individual form or the individual matter of substances are characterized primarily, the complex unity secondarily. Whole substances bear the ways they are derivatively. Blue characterizes the book because, primarily, blue is a way its material is.

With these remarks we can move on to the third element in our ideological toolbox: *qualitative identity*, or *qualitative sameness*. This is a relation whose relata are ways complex unities are.

In order to understand the notion of sameness needed here, we can follow Heil, who states that sameness is something which is "built in the ways substances are"[4]. Qualitative identity, or sameness, as the relation between ways substances are cannot be further analysed. More precisely, no further non-*formal* analysis is available. If ways substances are are qualitatively identical, then they are qualitatively identical full stop. The attribute "formal" is decisive here. Formal relations are introduced in the literature as relations which occur if and only if

4 Literally Heil (2003), p. 132, says: „Similarity is built in modes"; see also *ibid*., p. 157.

their relata occur.[5] That means that the occurrence of some relata is necessary and sufficient for the occurrence of the relations in question. Jonathan Lowe regards such relations as "entirely determined by their relata". They offer "no additions to reality". They are not entities in and of themselves, nor are they elements of beings (cf. Lowe (2006), p. 46).

That ways substances are are qualitatively identical, or the same, thus does not mean that we are committed to the existence of a relational entity between them. Nor does it mean that we have to look for another basis to ontologically ground them. Qualitative sameness is necessarily and sufficiently given with qualitatively identical ways substances are. The blueness that characterizes this book is qualitatively the same as the blueness that characterizes that book. Full stop.

To summarize: The possibility of giving a paraphrase cannot be decided absolutely or *simpliciter*, but only relative to a toolbox of ontological prerequisites. The ideas of *complex unities*, of *ways complex unities are*, and *qualitative identity of ways complex unities are* can be regarded as the prerequisites for a paraphrase of "there are natural kinds"-sentences. What we must do now is present the intended paraphrase.

3.3 The paraphrase of natural kinds-sentences

The main idea is that speaking about natural kinds, e.g. saying that spiders are such-and-such, does not commit us to sentences such as "there is something, namely the abstract natural kind spiderhood, and there are some other things which are abstract properties, which nomologically inhere in spiderhood";[6] rather, it commits us merely to sentences such as "there are some things, namely particular substances, spiders, which have an individual form; and there are other things, namely particular ways their individual forms are, which specifically characterize these forms". Putative (sentences about) natural kinds, e.g. spiders, are nothing more than (sentences about) particular substances with qualitatively identical individual forms, where these forms are characterized by the qualitati-

5 For more details concerning formal relations see my article „Existential Dependence and other Formal Relations" (Kanzian (2015)). We find the decisive hint for understanding formal relations in Kevin Mulligan's characterization of thin or internal relations. See Mulligan (1998), p. 344: „… a relation is internal with respect to objects a, b, c etc., just if, given a, b, c etc., the relation must hold between and of these objects."

6 The "nomological"-clause is taken from Lowe (2006), pp. 14 and 17.

vely same particular ways they are. This suggested reduction via paraphrase can be explicated and explained by appeal to our conceptual toolbox.

The idea is that every sentence about spiders is a sentence about particular spider-substances. As we have assumed, each spider is a complex unity, analyzable in its structure as consisting of a specific individual form and a material aspect. The individual form is characterized. Biologists may help provide an adequate list of characteristics, and these characteristics might be interpreted by ontologists as ways forms (or derivatively: whole spiders) are, spider-specifically. All of the ways the individual forms of two substances (e.g., of Susan and Sabrina) are may be qualitatively identical. Some authors would prefer to put this by saying that Susan and Sabrina are essentially similar. This way of speaking is harmless as long as these authors add: full stop. The qualitative identity of the ways upon which this essential similarity rests cannot be analyzed and requires neither a grounding instance nor a grounding entity.

Furthermore, for every substance, there is a determinate answer to the question whether it shares the ways its individual forms are with Susan and Sabrina, and to the question whether all of the ways their individual forms are are qualitatively identical with those of Susan and Sabrina. When we refer to spiders, we refer to nothing other than to Susan, Sabrina, and all of the other substances with the same ways their individual forms are. If whole substances with the same ways their individual forms are are essentially similar, we can conclude that natural kinds are nothing other than essentially similar substances. Our talk of spiders does not need abstract spiderhood and abstract spider-properties.

This will become even clearer when we try to interpret some examples of standard sentences about spiders. Consider for example the sentence asserting that spiders have typical characteristics, e.g. that they have eight legs. In order to interpret "Spiders have eight legs", we do not need an assumption such as: "There is the abstract kind, spiderhood, and there is the abstract property of having eight legs which nomologically inheres in spiderhood". What we really mean by "Spiders have eight legs" is: "There are … particular substances, spiders, and they have an individual form which is characterized specifically, and their having eight legs is due to one of these characterizing ways their respective forms are."

Another important case is given by the comparison of different (putative) natural kinds. As we have learned, not only spiders have eight legs, but so do other animals, e.g. octopi. So it is reasonable to assert: "Spiders and octopi have the same number of legs". Recall that our task is not to understand statements about the sameness of numbers, e.g. numbers of legs, but rather to understand spiders and octopi, while avoiding commitment to "There are abstract natural kinds"-sentences. The relevant suggestion is to say: "There are … particular substances which each have an individual form … that is characterized spider-specifically,

and there are … particular substances which have an individual form … that is characterized octopus-specifically, and all of them have eight legs due to one of the characterizing ways their forms are." This is the same as claiming that (all) substances with individual spider-forms are similar to (all) substances with individual octopus-forms, in the sense that their leg-number is fixed according to their individual form, which itself is one of the ways their respective forms are.

There is a price for admitting a plurality of similarities, but I see no problem in paying it. In addition to the abovementioned essential similarity (of e.g. all spiders), which amounts to the qualitative identity of all form-characterizing ways, we should also accept similarities that rely only on some form-characterizing ways. The crucial point is that none of these similarities amounts to an addition to reality, and all can be derived conceptually from the basic tie, namely qualitative identity. Positing a plurality of similarities helps us understand comparisons of spiders, and of spiders and octopi, without adding any complications to our ideology or increasing the number of required categories.

The last spider-statements we must take into account are expressions that subordinate spiders under a more general kind. Take for instance: "Spiders are animals". The abstracta-avoiding interpretation of this sentence requires a further distinction concerning our individual forms, more precisely concerning the characterization of individual forms by the ways these forms are. The distinction arises between the manners in which the characterization might be done. A characterization could be done in a (detailed) *specific* or in a (non-specific) general or *generic* manner. The more specifically we characterize forms, the more ways those forms are we appeal to. The most specific characterization of a form considers all of the ways that form is. It may also be called determinate or complete characterization. With this distinction in hand, we can reconstruct the classical distinction between the most specific kinds, or *species infimae*, and more general kinds, or *genera*. The distinction is based on the number and the extent of the form-characterizing ways that we take into account. That a substance belongs to a *species infima F* means no more than that its individual form is characterized in a most determinate or complete manner by *F*-specific ways. That a substance belongs to a more general kind, or to a genus *G*, means that its individual form is characterized in a less determinate or incomplete manner by *G*-relative ways, where *F* may be regarded as a species within genus *G* just in case that the *G*-ways of characterizing forms are a subclass of the *F*-ways forms are. For example, that spiders are animals means that substances belonging to the specific kind spider also belong to the more general genus *animal*. More precisely: there are particular substances, and these substances have an individual form which is characterized as completely spider-specific, and the totality of ways their individual forms are has as a proper subclass the animal-relative ways individual forms are.

3.4 Is our paraphrase successful?

As a path to moderate nominalism, does our avoidance of abstract natural kinds succeed? A minimal success-condition is that the proposed reduction of natural kinds meet the adequacy-criteria for paraphrases given above. To recall, the first criterion is that the original sentence seem to imply the existence of so-and-so's, whereas the paraphrase does not imply the existence of so-and-so's. In the preceding sections of this paper I have aimed to show that paradigmatic spider-sentences, which could seemingly be interpreted as "there are natural kinds"-sentences, are actually sentences about particular organisms. If the reader already disagrees with my suggestions, repetition will not convince her.

The second criterion is that the paraphrase could be used for all of the same purposes for which the original is used in everyday life. The proposed interpretations of spider-statements sound more complicated than our everyday talk about spiders, of course. But this does not tell against this criterion. The proposed interpretations are complicated because they try to make the ontological prerequisites transparent. But in fact they come close to the everyday use of spider sentences. And even more importantly, they come at least as close to this usage as an abstracta-friendly interpretation would – and the latter interpretation can be expected, at least in its explicit form, to be even more complicated than my nominalist one. Or, more carefully: A nominalist paraphrase of original abstracta-friendly interpretations of spider statements does not distance us from everyday usage.

This matter doubtless requires further discussion. Hence this criterion should be considered in the context of the other criteria, namely that the theoretical basis of our paraphrases should be consistent and economical. There is no reason to think the proposed are inconsistent; moreover, we may be optimistic, for they are so sparse that it would be hard for them not to be consistent. Complex unity, characterization, and qualitative identity are about as straightforward as we can get. It may be that some applications of these criteria are problematic and that others are difficult to explain. But at the level of our conceptual toolbox, there is every reason to think that the consistence-criterion is met. The same holds for the economy criterion, particularly in comparison with the abstracta-friendly alternative, which not only leads to a multiplication of entities, but, probably also yields a serious proliferation of basic concepts (inherence, instantiation, and so forth).

Last but not least, I want to discuss the requirement that a paraphrase should have the same implications for the relevant theoretical contexts and scientific practice as the original sentence has. The supposed advantages of abstract natural kinds are as follows. a) They explain the similarities between particular substances (e.g., between spiders) in accordance with David Armstrong's "one over many" principle: there is one abstract spiderhood which is instantiated in many

particular organisms. The second supposed advantage is b) that abstract natural kinds provide a clear and determinate condition for being (e.g.) a spider, that is, to instantiate spiderhood. Finally, c) that they offer an ontological basis for biological taxonomy: there is something which is a kind and stands in complex relations to other kinds. Our examples of spider-sentences are motivated by the intention of showing that a nominalist paraphrase also enjoys all of these theoretical merits. In applying our basic concepts, we can explain the similarity between particular spiders, demarcate spiders from non-spiders, and last but not least, suggest an alternative basis for biological taxonomy.

Space does not permit discussing every detail of implications of my theory of paraphrasing of natural kinds-sentences. So I will finish with some remarks on other candidates for abstract entityhood: universal abstract properties, and mathematical objects such as numbers, sets and classes.

4 Some final words on properties and mathematical objects

In section 3.1 we assumed that any attempt to reduce abstract properties and mathematical objects would be futile. This holds particularly for the former, because, as we saw, a successful paraphrase of "there are abstract properties"-sentences can be deduced immediately from a successful paraphrase of "there are abstract natural kinds"-sentences.

This result comes from the conceptual toolbox that we employed for our reduction of abstract natural kinds. The claim that "complexity", "characterization by ways the complexity is", "qualitative identity" are primitive notions in fact leads to a particularistic understanding of properties – or of modes, as we may follow Heil and Lowe in calling them. And it comes from the assumption that the reduction of universal abstract kinds makes universal abstract properties theoretically superfluous.

There are no universal properties without universal kinds. For an explanation of the connection between universal kinds and universal properties, we may refer to Jonathan Lowe and his *Four Categorial Ontology*. According to Lowe, the theoretical function of universal properties is to explain law-like dispositions of universal kinds.[7] The universal property of being a fly-eater, for instance, bestows on the universal kind of spiders a specific universal disposition which can be ex-

7 Cf. note 6.

pressed in terms of a biological law-like regularity: Spiders eat flies. But if there are no universal kinds, we have no reason to hold onto universal properties. If we have good reasons to deny the bearers, universal kinds (where such reasons would be given for instance by a successful paraphrase of universal kinds-sentences), then we can also get rid of the entities that (seemingly) inhere in them: universal properties.

But aren't there other theoretical reasons to accept universal properties? – I do not think so. However, rather than evaluating possible candidates for such reasons, we may assert the neglecting universal properties does not force us to entirely give up on properties and all of the theoretical advantages that they bring (such as the possibility of providing an ontological basis for the biological truism that spiders are fly-eaters). Particular modes of particular spiders do the job, and indeed they do it much better than universals – especially if we employ an ontological theory that has recourse to our proposed ideological toolbox.

Let us end with a remark on the futility of reducing mathematical objects. As we saw, this task is allegedly futile simply because the criteria for success do not seem to apply given the ideological toolbox at our disposal, at least in the context of a Quine/van Inwagen-metaontology. But aren't there other paths from moderate nominalism to extreme nominalism, that is, from a nominalism that neglects universal kinds and properties to one that avoids commitment to all abstract entities? This question seems to be synonymous with the questions: Isn't there an alternative to Quine's and van Inwagen's metaontology? Aren't there other success-criteria for paraphrases? Is there a more adequate ideology? Again, I do not think so. This is why we should prefer being moderate and more or less van Inwagenian.

Bibliography

Berto, F. and Plebani, M. (2015), *Ontology and Metaontology. A Contemporary Guide*, London: Bloomsbury.

Heil, J. (2003), *From an Ontological Point of View*, Oxford: Clarendon Press.

Kanzian, Ch. (2015), "Existential Dependence and other Formal Relations", in *God, Truth, And Other Enigmas*, edited by M. Szatkowski, Berlin/Munich/Boston: De Gruyter, 183–196.

Lowe, E. J. (2006), *The Four-Categorial Ontology*, Oxford: Clarendon Press.

Mulligan, K. (1998), "Relations – Through Thick and Thin", in *Erkenntnis*: 48, 325–353, Extra-Volume: Analytical Ontology, edited by Ch. Kanzian and E. Runggaldier.

Rudder-Baker, L. (2007), *The Metaphysics of Everyday Life*, Cambridge: Cambridge University Press.

Van Inwagen, P. (2014), *Existence. Essays in Ontology*, Cambridge: Cambridge University Press.

von Solodkoff, T. (2014), "Paraphrase Strategies in Metaphysics", in *Philosophy Compass*: 9(8), 570–582.

Joanna Odrowąż-Sypniewska

The Problem of the Many: Supervaluation, Rough Sets and Faultless Disagreement

Abstract: In the paper I make three comments concerning the existing solutions to the problem of the many: supervaluationism, fuzzy sets and Lewis's combined solution consisting of supervaluation and almost-identity. First, I try to defend supervaluationism from the charge that the precisifications it postulates are not admissible, because they do not preserve penumbral connections and clear cases. I argue that two types of vagueness should be distinguished and that the requirements that are imposed on precisifications postulated in one case do not apply in the other case. Next, concerning the solution advocating fuzzy sets I suggest that rough sets rather than fuzzy sets might be regarded as a good way of looking at composite objects. And finally, I point out to a surprising consequence of Lewis's combined solution. Namely, in my view Lewis's solution leads to the acceptance of faultless disagreements concerning the number of objects of the same type.

1 Introduction: The problem and its solutions

The problem of the many is due to Peter Unger (1980). David Lewis in his classic paper "Many but Almost One" formulated it as follows:

> Think of a cloud – just one cloud, and around it clear blue sky. Seen from the ground, the cloud may seem to have a sharp boundary. Not so. The cloud is a swarm of water droplets. At the outskirts of the cloud the density of the droplets falls off. Eventually they are so few and far between that we may hesitate to say that the outlying droplets are still part of the cloud at all; perhaps we might better say only that they are near the cloud. But the transition is gradual. Many surfaces are equally good candidates to be the boundary of the cloud. Therefore many aggregates of droplets, some more inclusive and some less inclusive (and some inclusive in different ways than others), are equally good candidates to be the cloud. Since they have equal claim, how can we say that the cloud is one of these aggregates rather than another? But if all of them count as clouds, then we have many clouds rather than one. And if none of them count, each one being ruled out because of the competition from the others, then we have no cloud. (Lewis (1993), p. 164).[1]

1 Quine's formulation of the problem is much shorter: "Who can aspire to a precise intermolecular demarcation of a desk? Countless minutely divergent aggregates have equal claims to being my desk." (Quine (1985), p. 167). Unger suggests that a better name for this problem might be "the problem of the many or the none".

https://doi.org/10.1515/9783110664812-019

Hence, we have arrived at a paradox: in the face of the existence of distinct equally good candidates for being the cloud, we seem to be compelled to conclude either that there are many clouds or that there are none; and both these conclusions contradict the common view that there is just one cloud.

Lewis also modifies P. T. Geach's paradox of 1001 cats (Geach (1980)) so as it becomes an instance of the problem of the many. The paradox concerns Tibbles the cat sitting on the mat. Lewis asks us to imagine that Tibbles is shedding. Some of her hairs become gradually looser and there will be hairs which are neither determinately parts of Tibbles nor determinately not parts of her. Let us assume that it is hairs h_1, h_2, ..., h_{1000} which are such questionable parts. Let c include all these hairs; c_1 include all of the hairs except for h_1, c_2 include all of the hairs except for h_2 and similarly for other 998 cs. Now, all the cs (c as well as c_1, c_2, ..., c_{1000}) have an equal claim to be the cat. So, we face the same problem as we have encountered in the case of the cloud.

The problem of the many arises for all materially composite objects: mountains, desks, people as well as clouds and cats. One can state it concisely thus:

> For many composite objects X there are objects y_0, ..., y_n such that it is indeterminate whether they are part of X or not.
> For each such object y_i there is an object Z_i which consists of y_i and all the unquestionable parts of X.
> Z_i are so similar to X, that they should count as X (i.e. there is no obvious reason not to count them as X).[2]

Faced with these premises we have two contrasting intuitions (Lopez de Sa (2008, 2014)). The grounding intuition compels us to say that since there are many things which have equal claim to be an X, we must conclude that there are many Xs, whereas the counting intuition tells us that there must be only one X. The solutions to the problem of the many can be divided into those that reject the grounding intuition in favour of the counting intuition (so called 'solutions by disqualification') and those that discard the counting intuition (and has been called 'egalitarian solutions'). Lewis in his paper mentions two types of the solutions by disqualification: those that argue that none of the many are clouds (i.e. none of the Zs is an X) and those that claim that one of the many is the cat (i.e. one of the Zs is the X). Among the former are the constituters solution and the vagueness-in-the-world solution. According to the constituters solution constitution is not identity and although the many are cat-constituters, none of them is a cat. Cats are constituted by parcels of matter but are not identical to them. The advocates of this

2 Odrowąż-Sypniewska (2005), p. 122.

solution have to introduce a distinction between parcels of matter and things and argue that although the matter constitutes the thing (perhaps for its entire history), it is not identical with it. According to the solution postulating vagueness in the world the many are cat-precisifications, but none of them is a cat. The many are not cats, for they are all precise, while cats are vague. This solution places vagueness in the world and assumes that there are vague objects such as cats, in addition to the precise ones, such as cat-precisifications. The precisifications are not identical with the cat, nor do they constitute it. They are a separate kind of objects existing alongside vague objects.[3] The supervaluationist solution belongs to the other type of solutions by disqualification. According to it only one of the many precisifications is the cat. "Cat" is vague and refers indeterminately to indefinitely many admissible cat-precisifications. Although there is no unique correct interpretation of the word "cat", there are many admissible interpretations, each one picking out one of the cat-precisifications to be the referent of "the cat". Hence, on each admissible interpretation just one of the many is the cat. Admittedly, on different sharpening a different cat-precisification will count as the cat, but this is not worrying, since always there will be only one cat. This is essentially a traditional supervaluationist solution to the sorites paradox. It does not posit the existence of vague objects and locates vagueness exclusively in the language.

Egalitarian solutions give priority to the grounding intuition and argue that the many are cats. However, they add that the cats are not many (i.e. the many *Z*s are the *X*s, but the *X*s are not many). Among such solutions Lewis mentions the relative-identity solution and the almost-identity solution. The former is a solution put forward by Geach, who believed in relative identity: two objects may be the same *F*, without being the same *G*. Thus, the many *c*s are both lumps of feline tissue and cats. They are different lumps of feline tissue, but they are all the same cat. The relation "is the same cat as" is a relative identity relation; it expresses "only a certain equivalence relation, not an absolute identity restricted to cats" (Geach (1980), p. 216). "Is the same cat" and "is the same lump of feline tissue" are different partial indiscernibility relations. When we count, we count by relative identity, not by absolute identity. Thus, although there are many lumps (counted by the "is the same lump of feline tissue as" relation), there is only one cat (counted by the "is the same cat as" relation).

According to the almost-identity solution identity is the relation of complete overlap. The many *c*s do not overlap completely, but they overlap *almost completely*. Hence, they are almost-identical. If we count by almost-identity relation, we will conclude that there is only one cat present on the mat.

3 See however Merlo (2017) (see below).

> The cats are many, but almost one. By a blameless approximation, we may simply say that there is one cat on the mat. (Lewis (1993), p. 178)

Lewis suggests that we should accept two solutions: supervaluationism and almost-identity. He argues that they work best when combined. Supervaluationism on its own works too well: if on each admissible precisification there is only one cat, then the problem of the many should not arise. On the other hand, the almost-identity solution alone does not apply to all cases. It doesn't work well for the case of Fred's house: it can either be taken as including the garage or as not including the garage.[4] Both interpretations have equal right to be regarded as Fred's house but we do not want to say that he has two houses. The almost-identity solution does not help here, because the house without the garage and the house including the garage do not almost completely overlap. However, supervaluationism provides an easy answer: "house" is semantically indeterminate and can be precisified in such a way that it either does or does not include the garage. In each precisification Fred has only one house. Another reason why almost-identity solution needs supervaluationism is the intuitive truth-value of statements like "The cat has the hair h_{17}" (see Lewis (1993), p. 181). Intuitively this statement is gappy, and this is precisely the verdict that supervaluationism delivers (since the statement is neither true not false on all admissible precisifications, it is neither supertrue nor superfalse).

One might also try to divide the solutions into those that regard the problem of the many as a problem of metaphysics and those that consider it a semantical problem. Clearly supervalutionism belongs to the latter group. It assumes that the problem arises because of semantic indecision. We have never made the decision which of the many candidates is the cat:

> Which one deserves the name 'cat' is up to us. If we decline to settle the question, nothing else will settle it for us. (Lewis (1993), p. 172)

Also the almost-identity solution regards the problem of the many as semantic in nature. Robert Williams and Dan Lopez de Sa notice that supervaluationism and almost identity agree on what there is, namely – in the case of mountains – "billions of mountainy agglomerations of rock" (Williams (2006), p. 418). They disagree only as to whether all mountain-candidates should be called 'mountains' or not. Supervaluationism claims that there is only one mountain, whereas almost

4 The example comes from Johnston (1992), p. 101. Lewis calls it "the problem of the two" (Lewis (1993), p. 180).

identity view has it that there are many mountains, but they may be counted as one.

At first glance it appears that the solution positing cat-constituters and the one appealing to vagueness in the world are metaphysical solutions. However, the matter is not that simple. For instance, Mark Johnston, who is one of the defenders of the view that constitution is not identity, acknowledges that there is "no credible metaphysical extra" (Johnston (1992), p. 103), which would justify the distinction between an object and its constituting matter. He claims however, that we should not draw the conclusion that there is no justifiable distinction to be made at all. It is true that the distinction is not a metaphysically grounded one, but we can justify its existence by turning to our practice. It is a fact that we do distinguish between an object *F* and its *F*-shaped constituting matter. A practice-dependent justification consists in first asserting that we do observe the distinction in our practice, and then giving "an internal and pragmatic justification of this in terms of how practice which marks this distinction serves our purposes" (Johnston (1992), p. 102). Johnston concludes that:

> [O]ur practice and the distinction it embodies is acceptable as it stands and what is bogus is the conception of justifying our practice which requires that, for the distinction to be justified, the difference between an *F* and its constituting matter must be a deep metaphysical difference secured by an extra ingredient of the *F*. (Johnston (1992), p. 103).

Similarly, Giovanni Merlo, who is an advocate of vagueness in the world, in a recent paper argues that there are vague objects, but they are nothing 'over and above' their many precise counterparts; "facts involving the vague object (…) *require no more of the world* than facts involving the object's precise counterparts" (Merlo (2017), p. 2651). If one sides with Johnston and Merlo, then neither of their respective solutions qualifies as truly metaphysical.

In what follows I would like to make three comments concerning the existing solutions. First, I'll try to defend supervaluationism from the charge that the precisifications it postulates are not admissible. Next, concerning the solution advocating fuzzy sets I'll suggest that rough sets rather than fuzzy sets might be regarded as a good way of looking at composite objects. And finally, I'll point out to a surprising consequence of Lewis's combined solution.

2 Are there any admissible precisifications?

Supervaluationism is best known for its treatment of vague gradable predicates such as "red", "tall" or "rich". It is commonly assumed that such predicates have

no definite boundaries.[5] Even if we fix a relevant comparison class (lets say Polish male adults), there will be people belonging to that class who will be neither clearly tall not clearly not tall. Thus, we may say that there are clear cases of being tall Polish adult male, clear cases of not being tall Polish adult male and borderline cases. According to supervaluationism we may precisify "tall" in many ways. Admissible precisifications will be those which sharpen the predicate in such a way as to introduce the cut-off point somewhere in the region of borderline cases (i.e. within the so-called penumbra). Such precisifications have to preserve clear cases and cannot introduce the boundary in such a way that a person that before precisification counted as clearly tall becomes not tall afterwards. Admissible precisifications must also preserve penumbral connections. For instance, if Bill is 177 cm tall, Frank is 178 cm tall and both are borderline cases of being tall, then any admissible precisification that makes Bill tall must make Frank tall as well.

Several philosophers have raised a worry that supervaluationism cannot be an adequate solution to the problem of the many because precisifications used in the solution are not admissible: they do not preserve penumbral connections and do not preserve clear cases.[6]

2.1 Penumbral connections

Precisifications of vague gradable predicates such as "tall" precisify the property of being tall in such a way that every object within the relevant class either clearly falls into to extension of "tall" or clearly belongs to the anti-extension. Those who advocate supervaluationism as a solution to the problem of the many have to say that it is predicates such as "is a cat", "is a coin" and "is a mountain" (and proper names such as "Tibbles") that are being precisified. However, in this case it is not the properties of being a cat, a coin or a mountain that are being made precise. Precisifications do not precisify the property of being a coin or a mountain in such a way that only one object has it. It appears that on each precisification an arbitrarily chosen candidate is elected to be a cat, a coin or a mountain. Neil McKinnon (2002) notices that being a coin or a mountain is not a fundamental property: an object is a coin or a mountain in virtue of having some more primitive, fundamental properties. Therefore only principled precisifications which preserve penumbral connections should be admissible. He argues that a principle like (NAD) must hold:

5 For a different view see e.g. Williamson (1994).

6 See e.g. Lopez de Sa (2014).

> (NAD):
> For any coin and non-coin, there is a principled difference between them which forms the basis for one's being a coin and the other's being a non-coin (McKinnon (2002), p. 333).

Arbitrary precisifications are in clear conflict with (NAD), for they severe the connection between coinhood and those fundamental properties on which being a coin depends. In reply Weatherson notices that there is no principled intrinsic difference between a coin and a token, or a coin and a medal (Weatherson (2003), pp. 495–496), so if one wants to defend (NAD) one must allow principled *extrinsic* differences. And if we admit properties like "being created with the intent of being used in a coin-like way" (Weatherson (2003), p. 496), we can insist that it is this property that is being precisified on each sharpening. However, Weatherson admits that his solution works only for artefacts and cannot be extended, for instance, to mountains.[7] Precisifications of mountains seem to be arbitrary and as a consequence we have to admit that "facts about mountains are not grounded in mountainhood-free facts" (Sattig (2013), p. 214).

2.2 Clear cases

According to the supervaluationist treatment statements that come out true in all admissible precisifications are supertrue, statements that come out false in all admissible precisifiactions are superfalse and those that are true in some precisifications but false in some others are devoid of truth value. The operator 'Definitely' ('Def') which expresses supertruth maybe introduced: 'Def *A*' is true if and only if '*A*' is true in all precisifications. Thus, if someone is a clear case of 'tall' and is tall in all precisifications, he is definitely tall, whereas someone who is a borderline case of 'tall' is neither definitely tall nor definitely not tall.

7 He notices that (NAD) does not hold for mountains either. (Weatherson (2003), p. 497). Lopez de Sa notices that (NAD) is subject to the sorites paradox, so it should not be regarded as a penumbral truth but rather as "an intuitively appealing but ultimately rejectable soritical principle" (Lopez de Sa (2014), p. 1109). According to him a better candidate for penumbral truth is Unger's Principle of Minute Differences. He acknowledges that supervaluationists might feel compelled to reject PMD, just as they reject the sorites premise, but complains that they do not explain away the grounding intuition underlying PMD (see Lopez de Sa (2014), p. 1109). It seems to me however, that the reason why supervaluationists do not explain away the grounding intuition is that they fully acknowledge its force (after all each cat-candidate gets to be the cat on some precisification). They just think that the counting intuition, according to which the cat is just one, is even more powerful and should be given precedence.

When we apply supervaluationist treatment to the problem of the many we will see immediately that nothing is such that it is definitely a cat or a mountain. Since on each precisification a different cat-candidate counts as cat, there is no such thing that is a cat on all precisifications, hence there are no clear cases of "cat". It appears that all cat-candidates have to be regarded as borderline cases. This would really be an unwelcome result and no wonder that Lopez de Sa complains:

> There seems to be something deeply disturbing about the thought that there is no relevant difference between the individuals at the clear end and in the middle ground of the sorites series from cats to pet-robots: all of them are, according to the solution, merely borderline cases with respect to 'is a cat'. (Lopez de Sa (2014), p. 1110)

Let us focus on sorites series starting with a paradigm cat and ending with a pet-robot, such that each step in the series differs from the previous one only in that one tiny part of it has been exchanged for a mechanical part. Then, we would want to say that the first steps in the series are clear cases of cats, while last steps in the series are clear cases of robots. The problem is that there seem to be no point of which we would be willing to say that it separates cats from robots. There seem to be no clear boundary between cats and robots, and yet cats are at the one end of the series and robots at the other end. It is widely accepted that supervaluationism comes to help here. Supervaluationism applied to such a series makes it possible to argue that there is a boundary between cats and robots somewhere in the series without forcing us to fix this boundary at a particular place (for on any admissible sharpening there is such a boundary, but it shifts and is in a different place on each sharpening). However, if we agree that supervaluationism is also the best solution to the problem of the many we seem to face a serious problem. For what we have at the beginning of the series are not clear cases of cats but partially overlapping cat-candidates none of which is a cat on all sharpenings of "cat". Thus even the alleged clear cases turn out to be borderline cases after all and instead of a sorites series beginning with cats and ending with robots we have a series of borderline-cats. This indeed seems disconcerting.

2.3 Reply

I'd like to argue that the situation is not as bleak for supervaluationism as it might seem. Those authors who complain that precisifications to which defenders of supervaluationism appeal in their solution to the problem of the many do not preserve clear cases or penumbral connections, confuse two types of vagueness that are in play. This is most clear in a paper of Lopez de Sa "Is the problem of the many

a problem in metaphysics?". In this paper the author starts with a discussion concerning Kilimanjaro, which is a paradigmatic mountain and atom Sparky, which is neither determinately part of Kilimanjaro nor determinately not part of it. We may call Kilimanjaro together with Sparky "Kilimanjaro(+)" and Kilimanjaro without Sparky "Kilimanjaro(-)". After discussing the vagueness of "Kilimanjaro" which is due to the fact that it indeterminately refers to Kilimanjaro(+) and Kilimanjaro(-), Lopez de Sa goes on to say:

> Not only is 'Kilimanjaro' vague, but 'is a mountain' is too: clearly one can imagine a sorites series going from the paradigmatic mountain Kilimanjaro to the hill Montjuic. (Lopez de Sa (2008), 748)

It is of course true that both "Kilimanjaro" and "is a mountain" are vague, but the possibility of creating a relevant sorites series is not a justification for the claim that "is a mountain" is vague in a way that Lopez de Sa has just discussed. One type of vagueness has to do with not having sharp spatial boundaries (and this is the kind that Lopez de Sa discusses in connection with Kilimanjaro) and another type has to do with having borderline cases (i.e. with it being possible to imagine a relevant sorites series). Those two types are independent: as it happens both "Kilimanjaro" and "is a mountain" are vague in both ways, but it is a contingent matter. We can imagine that people decide to precisify "mountain" in such a way that only a hill that is more than 2000 m high counts as mountain. If this definition where accepted then there would be no borderline cases between "is a mountain" and "is a hill" and yet the problem of the many would still arise for each mountain. Similarly, even if there were objects with absolutely precise spatial boundaries, they still could figure in a sorites series for a relevant predicate (such as "is a mountain") and that predicate would have borderline cases.

Such two types of vagueness have already been distinguished by Quine in "Word and Object":

> Commonly a general term true of physical objects will be vague in two ways: as to the several boundaries of all its objects and as to the inclusion or exclusion of marginal objects. Thus take the general term 'mountain': it is vague on the score of how much terrain to reckon into each of the indisputable mountains, and it is vague on the score of what lesser eminences to count as mountains at all. (Quine (1960), p. 126)

It is the first type of vagueness that gives rise to the indeterminacy of count[8] and to the problem of the many. The second type leads to the sorites paradox. Once we notice that two kinds of vagueness are involved we may distinguish two ways in which an object might be a clear case of a predicate. On the one hand, Tibbles the cat is definitely a cat when we compare her with pet-robots, but on the other hand there is no such object of which we may say that it definitely is Tibbles. Again, Quine diagnoses the situation ingeniously:

> The extension of the term 'desk' is conventionally thought of as the class of its denotata, thought of as physical objects. Realistically we may recognize rather an *extension family*, as I shall call it. It is a family of vaguely delimited classes, each class being comprised of nested physical objects any of which would pass indifferently for one and the same desk. (Quine (1985), p. 168)

Thus, when we are faced with a sorites series and we try to find a solution, we choose to forget that instead of one object at each step we have indefinitely many of them. Once we concentrate on objects at each step separately and pay attention to their spatial boundaries we will realize that our former considerations were based on a blameless simplification. Preserving penumbral connections and clear cases is crucial for the sorites kind of vagueness but is not important for the fuzzy-boundaries kind.[9] The claim that none of the cat-candidates is determinately a cat is compatible with the claim there are objects (or rather classes of objects in Quine's terminology) that are clear cases of the predicate "is a cat". The requirements that are imposed on admissible precisifications of gradable scalar predicates like "tall" and "rich" need not – and as a matter of fact do not – apply to admissible precisifications of predicates like "is a mountain" or "is a coin"[10].

2.4 Interlude: Is multiple reference better than indeterminate reference?

In a recent paper Merlo proposes a new solution to the problem of the many and argues against supervaluationism. His main argument has to do with the truth-

8 "The first of the two ways in which 'mountain' is vague causes an indeterminacy of count: it is not clear when to declare a saddle to be in the middle of one mountain and when between two mountains. The issue makes all the difference between one mountain and two" (Quine (1960), p. 126).

9 See also footnote 7 above.

10 Obviously, there are other requirements that such precisifications must satisfy (e.g. any admissible precisification has to include all determinate parts of a cat/mountain).

values of some statements, which supervaluationism dictates but which are extremely counterintuitive. For instance, Merlo notices that according to supervaluationism all the three following statements are true:

(1) Kilimanjaro has no precise number of atoms.
(1*) There is no number *n* such that *n* is determinately the number of atoms of Kilimanjaro.
(1**) Kilimanjaro is something which has a precise number of atoms.

At first glance, it might seem that supervaluationists are compelled to argue – counterintuitively – that (1) if false, because on each precisification Kilimanjaro has a precise number of atoms. To avoid this conclusion they suggest that we should paraphrase (1) as (1*). (1*) is supertrue, because on each precisification of Kilimanjaro different *n* will be the number of atoms of Kilimanjaro. However, it appears that (1**) should also be true, because on each precisification it is true that Kilimanjaro is something which has a precise number of atoms. Thus, Merlo remarks:

> But if we adopt the proposed interpretation of "has precise number of atoms", I am right in believing Kilimanjaro to have no precise number of atoms (because (1) is equivalent to (1*) which is true) and wrong in believing it to be something which has no precise number of atoms (because (1**) is also true)). (Merlo (2017), p. 2649)[11]

In Merlo's opinion those problems are serious enough to abandon supervaluationism[12] and look for another – semantic – solution. For the lack of space, I'll not describe his proposed view in detail, but I'll concentrate on his main claim, which is that terms like "Kilimanjaro" are multiple terms, where *b* is a multiple term iff many things are each exactly denoted by *b* (Merlo (2017), p. 2654). The exact denotation of a term *b* is the best answer to the question "What does *b* denote?" Among non-singular terms[13] are plural terms like "the Beatles"[14] and

11 Another unwelcome result is that although "It is indeterminate whether Sparky is part of Kilimanjaro" comes out supertrue, "There is something such that it is indeterminate whether Sparky is part of it" is superfalse.

12 More precisely, he claims that although the difficulties that he discusses are not fatal, "what each of them shows, however, is that the idea of indeterminate reference can only get us so far when it comes to accommodating our ordinary talk of vague objects" (Merlo (2017), p. 2651). Later on he confesses that he finds attractive the idea that language contains both semantic multiplicity and semantic indecision (Merlo (2017), p. 2661).

13 A term is singular iff it denotes one and only one thing in each context in which it is used. (Merlo (2017), p. 2653)

14 Term *b* is a plural term iff many things are collectively exactly denoted by *b*. (Merlo (2017), p. 2654)

multiple terms like "Kilimanjaro". On Merlo's view "Kilimanjaro" multiply refers to precise objects K_1, K_2, K_3, ...[15] To obtain intuitive truth-value assignment another assumption is needed, namely that predicates ascribing vague features to Kilimanjaro and predicates ascribing precise features to K_1, K_2, K_3, ... are non-distributive, where a predicate F is distributive, if and only if it is analytic that F is true of b iff each of the things among b is F (Merlo (2017), p. 2653). This view allows one to say that while (1) and (1*) are true, (1**) is false. It also makes it possible to claim that all the following are true:

> Kilimanjaro has no precise number of atoms.
> Kilimanjaro has fuzzy boundaries.
> K_1 has a precise number of atoms.
> K_1 does not have fuzzy boundaries.

Those intuitive truth-value assignments are indeed an advantage of Merlo's view, but he does not seem to realize that his view forces one to accept an extremely counterintuitive claim, namely that vast majority of proper names are in fact non-singular terms. On this view all proper names of physical objects will be multiple terms, like "Kilimanjaro". When Merlo discusses the example of the plural term "The Beatles" he seems to be assuming that "Paul", "John", "George" and "Ringo" are like K_1, K_2, K_3, ... and not like "Kilimanjaro". However, the lesson that the problem of the many teaches us is that each of "Paul", "John", "George" and "Ringo" will multiply refer to precise objects J_1, J_2, J_3, ...; , and so on. People are presumably less fuzzy than mountains, but for them the problem of the many arises as well. Although Merlo could argue that just as Kilimanjaro is nothing 'over and above' K_1, K_2, K_3, ..., Paul is nothing 'over and above' P_1, P_2, P_3, ..., the consequence that effectively no proper names are singular terms seems too high a price to pay.[16] Merlo objected to supervaluationism on the grounds that it ascribes counterintuitive truth-values to some statements, but what he proposes instead is even more counterintuitive. We are left with a language in which there are virtually no singular terms, presumably save for proper names of abstract and non-composite objects, for which the problem of the many does not arise.

15 K_1, K_2, K_3, ... are precise mountain-like things located roughly where Kilimanjaro is located. See Merlo (2017), p. 2646.

16 Merlo claims that "Kilimanjaro is nothing 'over and above' K_1, K_2, K_3, ..., without being numerically identical to any of them, just as the Beatles are nothing 'over and above' their four members without being numerically identical to any of Paul, John, George and Ringo." (Merlo (2017), p. 2664)

3 Are fuzzy sets fuzzy enough?

3.1 Human beings as fuzzy sets of simples

In *Material Beings* Peter van Inwagen argues that behind the problem of the many applied to human beings, such as himself, there are the following three crucial assumptions (Van Inwagen (1990), p. 216):

(1) If he (e.g. Peter van Inwagen) exists, there is a man.
(2) In every situation of which we would ordinarily say that there is just one man, there are many sets of simples which are equally good candidates to compose a man.
(3) The members of each of these sets compose something.

Anyone who wants to hold on to these three assumptions and does not believe in spatio-temporal coincidence must find a suitable selection principle – i.e. in this case a principle which would allow to claim that one (and only one) of the huge number of overlapping composite objects is Peter van Inwagen. Either the others are men but are not van Inwagen or else they are not men, but merely men-candidates. Since the differences between the overlapping composite objects are negligible, the principle has to be "intolerably arbitrary" (Van Inwagen (1990), p. 216). Van Inwagen avoids this problem by rejecting the assumption (3). He claims that he exists, and that in the relevant situation there are many sets of simples whose members are suitably arranged to compose men, but he denies that the members of those sets compose anything. He denies that in such situations there are many things which are barely distinguishable and which are candidates to be him. On the contrary, he argues that although he is present in such cases, none of the other things present in those cases is even similar to him. In particular, according to him a collection consisting of simples that compose him minus one simple, does not compose anything at all.

According to van Inwagen the *x*s compose *y* if the activity of the *x*s constitutes a life, and constituting a life is obviously a vague notion. Being caught up in the life of an organism is not a precise condition, and it may be indeterminate of some simples whether they satisfy that condition or not. Therefore, composition and parthood are also vague notions: for some simples it can be vague whether they are parts of a living organism; it can be vague whether they compose something. In particular, there are simples such that it is vague whether they are parts of van Inwagen. Thus, he argues that no set is the set that contains just the simples that compose him. Sets have precise membership-conditions: each object must either be or not be a member of a given set. So the simples that compose him do not constitute a set. No matter which of the competing sets of simples we take, it will not

be the set of simples that compose van Inwagen. Hence, although there are many distinct sets of simples, there is no problem of the many: none of the many sets is van Inwagen. The simples that compose van Inwagen constitute a *fuzzy set*. Membership in a fuzzy set is a matter of degree. For each fuzzy set there are objects that definitely are members of that set, objects that definitely are not members of that set and objects that are neither definitely members nor definitely non-members of that set. These last objects are members of the fuzzy set only to a certain degree. It is usually assumed that there are as many degrees of membership as the real numbers from 0 to 1. Specifying a fuzzy set amounts to specifying for each object the degree to which it is a member of that set. Objects that are members to the degree 1 are definite members, objects that are members to the degree 0 are definite non-members, while objects that are members to the degree d, where $0 < d < 1$, are indefinite members of the fuzzy set in question.

The fuzzy set of simples whose members are parts of $x \stackrel{df}{=}$ The fuzzy set y of simples such that $\forall z$ a simple is a member of y to the degree z iff that simple is a part of x to the degree z (Van Inwagen (1990), p. 223).

Van Inwagen assumes that each simple (at any given time) is a part of x to some specifiable degree d, where $0 \leq d \leq 1$. It follows from this assumption and the above definition that (at any given time) the members of exactly one fuzzy set of simples will compose x, and in particular – van Inwagen. The degree of membership is the degree of parthood: the degree to which x is a member of F is the degree to which z is a part of y. Therefore the problem of the many disappears: there is only one fuzzy set whose members compose van Inwagen.

Thus, the problem of the many is not a problem anymore.[17] As van Inwagen himself notices, the solution hinges on the assumption that there is only one life present in the case in question. The crucial assumption for him is that every situation of which we usually say that it contains just one man, contains just one life. He does not try to justify that claim: he takes it to be a common sense assumption and argues that we have no reason whatsoever to believe that "there are very many more human lives than the census-takers say there are human beings" (Van Inwagen (1990), p. 227).[18]

17 Van Inwagen's solution applies only to living organisms, but – as it is well-known – according to him such organisms are the only composite objects.

18 This section appears in Odrowąż-Sypniewska (2000).

3.2 Rough sets

3.2.1 Objections to van Inwagen's solution

Terence Horgan in his review of *Material Beings* argues that van Inwagen does not succeed in solving the problem of the many, because there is no "unique *fuzzy* set F such that for each simple S, S is a member of F to degree n iff S is now a part of me to degree n" (Horgan (1993), p. 697). If we call "parthood distribution" an assignment of a real number to each of the simples that exist in the world, then there will be infinitely many similar parthood distributions, each of which has equal right to constitute van Inwagen. Since there is no "systematic, general, principle of selection that renders a unique one of these, as against any of the others, the actual composition of me now" (Horgan (1993), p. 698), van Inwagen's solution is committed to the following "unexplainable" (ibid.) fact:

> Fuzzy set F, rather than any of many other similar fuzzy sets, is the unique one whose graded membership relation coincides with the composition of organism O at time t.

And facts such as this violate the principle of the non-arbitrariness of composition, which Horgan takes to be fundamental and plausible and which van Inwagen himself presupposes when he insists on a general and systematic answer to the special composition question[19] (see Horgan (1993), p. 697).

Another problem for the fuzzy-set solution to the problem of the many is that it might be objected that fuzzy sets are not vague enough. Since each object can be ascribed a specified degree to which it is a member of a given fuzzy set, one might argue that vagueness disappears. For instance, there will be a clear boundary between objects that are members to the degree 1 and objects that are members to the degree lesser than 1. Hence, Michael Tye argued that "fuzzy logic, at least in its normal development, is not fuzzy enough to accommodate our ordinary intuitions about vagueness" (Tye (1996), p. 223).

19 The special composition question is: When is it true that there is a y such that the xs compose y? The answer van Inwagen gives and defends in chapters 1 -16 of *Material Beings* is the following: $\exists y$ the xs compose y iff the activity of the xs constitutes a life.
However, the problem of the many forces him to modify this answer. The refined version of the answer is this: "$\exists y$ the members of the fuzzy set of simples x compose y iff the activity of the members of x constitutes a life."

3.2.2 Orłowska's rough set theory

Faced with these objections one might argue that composite objects should be analysed as rough sets rather than as fuzzy sets. Such sets are determined by an indiscernibility relation and their boundaries cannot be precisely determined. The best one can do is to specify their upper and lower approximations, which determine the region within which the 'proper boundaries' are confined. The theory of rough sets has been developed by Zdzisław Pawlak (1983), while Ewa Orłowska (1985) has applied his theory to vague expressions.

Orłowska claims that we perceive properties of objects through an assignment of attributes and their values to objects:

> In general we are not able to distinguish single objects with respect to the admitted attributes, namely we identify those objects which have the same values for all these attributes. As a consequence we 'grasp' not single objects but some classes of them. In each class there are elements which cannot be distinguished one from the other by means of given attributes. (Orłowska (1985), p. 469)

Those classes are either disjoint or not. If not all the pairs of classes are disjoint then they determine a tolerance relation (which is reflexive and symmetric). One can define an approximation space as a formal counterpart of our perceptual ability of observation. An approximation space is any pair $\langle U, R\rangle$; where U – the universe of discourse, R – indiscernibility relation (binary relation on set U). R is a representation of our ability to perceive U and is either an equivalence or a tolerance relation. The classes of indistinguishable elements determined by R are called 'elementary sets'. In an approximation space $\mathfrak{S} = \langle U, R\rangle$; a subset X of U is definable in $\mathfrak{S}$ if it can be represented as a union of the elementary sets determined by R or if it is empty. (ibid.)

> Given a non-definable subset X of U, our observation restricted by R causes X to be perceived as a vague object. In fact, we perceive not the single set X but a family of those sets which cannot be distinguished from X. In other words, (...), we observe X with some tolerance. (Orłowska (1985), p. 469)

The limits of tolerance are determined by a pair of definable sets: an upper approximation $\overline{R}X$ of set X is the least definable subset of U containing X, whereas a lower approximation $\underline{R}X$ of set X is the greatest definable subset of U contained in X. (Orłowska (1985), pp. 469-470)

Orłowska considers the following example (Orłowska (1985), p. 471). Let U be a set of men and R a relation of having the same number of hairs. In each elementary set of this relation there are men who cannot be distinguished by the attribute "number of hairs". Let us assume that X is the set of bald men from U. Then,

the attribute "number of hairs" is not sufficient to define X in our approximation space.

> The lower approximation of X consists of those men who are definitely bald with respect to the number of hairs, and the upper approximation of X consists of men who are possibly bald with respect to this attribute. (Orłowska (1985), p. 471)

If our perception of elements of U is limited to certain attributes, we may not be able to classify them into those that have a certain property P and those that do not. "In general, we are not able to decide for each element of U whether it belongs to X_P [i.e. to the subset of all elements from U for which P holds] or not, we can decide it up to relation R." (Orłowska (1985), p. 471) Hence, we approximate the extension from below and from above using classes of indiscernibility: an element u is a positive instance of P iff u belongs to $\underline{R}X_P$, whereas an element u is a negative instance of P if and only if u belongs to U - $\overline{R}X_P$. (ibid.)

3.2.3 Composite objects as rough sets

At the beginning of the paper in which Orłowska applies rough sets theory to vague expressions she suggests that she regards vagueness as a semantic phenomenon. She mentions "vagueness understood as a deficiency of meaning" (Orłowska (1985), p. 465) and regards the lack of boundaries as "semantically a deep phenomenon" (Orłowska (1985), p. 466). However, at the same time she writes that "we shall consider epistemological vagueness of predicates consisting in unavailability of total information about the universe" (Orłowska (1985), pp. 465-466). It is this lack of information which is supposed to cause the indeterminacy of the meaning of predicates. The lack of sharp boundaries is an alleged consequence of "the fact that we describe a continuous world in observational terms, and usually there are entities or properties which cannot be grasped by observation" (Orłowska (1985), p. 466). These last remarks suggest that Orłowska treats vagueness as an epistemological rather than semantic phenomenon, due to the lack of information about the world and imprecision of our observational abilities. I once argued that regarding vagueness of predicates such as "bald" or "tall" as epistemic in nature is a drawback of Orłowska's approach (see Odrowąż-Sypniewska (2000)). I insisted in particular that vagueness of such predicates is not due to our ignorance and that even if we had all the information concerning the number of hair of the people in a relevant group we would still not be able to divide them into bald and not bald ones. Knowing the exact number of hair would not make vagueness of "bald" disappear. I'd like to point out however that Orłowska's re-

marks fit much better the fuzzy-boundary type of vagueness related to the problem of the many. In fact, it might be argued that rough set theory could be extended in such a way as to serve as a solution to that problem. We may take U to be set of simples (or molecules, or rocks) and R to be the relation of being part of the same cat or mountain. R is an indiscernibility relation holding between those simples or molecules that are part of the same composite object, while X is the relevant composite object in question (e.g. a cat or a mountain). The relation R will have to be defined in terms of more primitive, fundamental properties.[20] Anna Wójtowicz in a paper "Some philosophical aspects of indiscernibility" shows how to extend this account to binary predicates.[21] Thus, it appears that if we regard relations such as being part of the same cat or the same mountain as indiscernibility relations and if we manage to determine more basic properties and relations in which those indiscernibility relations are grounded then we might be able to think about composite objects as rough sets. If cats and mountains are rough sets of their parts rather than fuzzy sets then the objections raised by Horgan and Tye do not apply: there is no problem with there being no unique parthood distribution that constitutes the composition of the organism O and there is no danger that vagueness will go away (for we do not assume any parthood distribution).

4 Faultless disagreement concerning the number of objects present

Recently in the philosophy of language literature there has been quite a lot of discussion concerning so-called faultless disagreement. Max Kölbel defines faultless disagreement as follows:

20 Orłowska gives the example of the set of people suffering from leukaemia, which can be characterized by relation R, which will be grounded in properties like having certain values of blood parameters and certain results of myeloblast test. See Orłowska (1985), p. 473.

21 Wójtowicz gives the example of people suffering from a certain illness A and shows that suffering from A can be stated in terms of properties (like having a certain blood group) and relations (like being a sexual partner). She uses Quine's partial notion of indiscernibility and assumes that the relations in question are symmetrical: "$x\ ind_P\ yiff \forall P^1[P^1(x) \leftrightarrow P^1(y)] \wedge \forall P^2\{[(\forall v, z[P^2(v,z) \leftrightarrow P^2(z,v)]) \rightarrow P^2(x,y)\} \wedge \{\neg\forall v, z[P^2(v,z) \leftrightarrow P^2(z,v)]\} \rightarrow (P^2(x,y) \leftrightarrow P^2(y,x))] \vee (x = y)$" (Wójtowicz (1998), p. 388).

> A faultless disagreement is a situation where there is a thinker A, a thinker B, and a proposition (content of judgement) p, such that:
> (a) A believes (judges) that p and B believes (judges) that not-p
> (b) Neither A nor B has made a mistake (is at fault) (Kölbel (2003), p. 54);

whereas Chris Kennedy says that faultless disagreement occurs

> ... if a sentence and its negation can be used by competent speakers to contradict each other, but in a way that (from external point of view) appears to be consistent with both speakers saying something true. (Kennedy (2013))

It is usually claimed that disagreements of this sort appear in conversations concerning ethical or aesthetical matters. The most common examples involve predicates of personal taste, such as "fun" or "tasty". For instance, Isidora Stojanovic considers the following exchange in which Tarek and Inma comment upon soybean ice-cream they have just tasted:

1. *Tarek*: This is delicious.
2. *Inma*: That's not true. This isn't delicious at all.

Stojanovic remarks:

> On the one hand, we are inclined to say that Tarek and Inma disagree. But on the other hand, we are also inclined to say that Tarek and Inma may both be right, and that their seemingly contradictory utterances may be true together. (Stojanovic (2007), p. 2)

It has also been argued that such disagreements may concern borderline cases of vague scalar predicates, such as "rich" or "tall" (see e.g. Kennedy). So for instance, if A says 'Philip is tall' and B says 'Philip is not tall', their disagreement may faultless if Philip is a borderline case of tallness. In such a case it seems that both A and B maybe right even though they utter what appears to be contradictory statements.[22]

The very existence of faultless disagreements is a contentious matter, since many people argue that in cases in which interlocutors appear to disagree faultlessly they are not really disagreeing but merely talking past each other. Even those who are in favour of there being such disagreements do not usually postulate their existence in other domains. Therefore it might come as a surprise that Lewis's favourite solution to the problem of the many seems to make possible faultless disagreements concerning the number of the objects of the same kind

22 Were Philip clearly tall, then obviously the disagreement would be genuine, not faultless and only A would be right.

present in a given situation. As we remember, Lewis proposes a combined solutions consisting of supervaluation and almost-identity. He argues that there are two kinds of intended interpretations: those that regard all cat-candidates as cats and those according to which the cat is just one. He writes:

> Context will favour one sort of interpretation or the other, though not every context will settle the matter. (Lewis (1993), p. 179)

Now, if indeed there are contexts which do not settle the matter and which do not favour one interpretation over the other, and if we imagine that in just such a context *A* and *B* say:

A: There is (exactly) one cat on the mat.
B: That's not true, there is more than one cat on the mat,

then they clearly disagree and moreover their disagreement appears to be faultless. Since the context does not favour any of the two interpretations, both seem to be equally correct. Thus, if we assume that *A* and *B* believe in what they say, i.e. *A* believes that *p*, *B* believes that not *p*, they both appear to be right and what they say appears to be true. Frege famously argued that you cannot count objects unless you have a relevant criterion of identity: to count red things you need to know what counts as the same red thing (See Geach (1980), p. 177). However, on Lewis's view knowing the criterion of identity does not help, since two ways of counting (by identity and by almost-identity) may be equally right in a certain context. Thus, the resulting disagreement will be faultless.

Bibliography

Geach, P. T. (1980), *Reference and Generality*, Cornell University Press.
Horgan, T. (1993), "On What There Isn't", in *Philosophy and Phenomenological Research*: 53(3), 693–700.
Johnston, M. (1992), "Constitution is not identity", in *Mind*: 101, 89–105.
Kennedy, Ch. (2013), "Two Sources of Subjectivity", in *Inquiry*: 56, 258–277.
Kölbel, M. (2003), "Faultless Disagreement", in *Proceedings of the Aristotelian Society*: 104, 53–73.
Lewis, D. (1993), "Many but Almost One", in D. Lewis, *Papers in Metaphysics and Epistemology*, Cambridge University Press.
Lopez de Sa, D. (2008), "Is the Problem of the Many a Problem in Metaphysics?", in *Nous*: 42, 746–752.
Lopez de Sa, D. (2014), "Lewis vs. Lewis on the Problem of the Many", in *Synthese*: 191, 1105–1117.

Lowe, E. J. (1995), "The Problem of the Many and the Vagueness of Constitution", in *Analysis*: 55, 179–182.
McKinnon, N. (2002), "Supervaluations and the Problem of The Many", in *The Philosophical Quarterly*: 52, 320–339.
Merlo, G. (2017), "Multiple Reference and Vague Objects", in *Synthese*: 194, 2645–2666.
Noonan, H. (1993), "Constitution is identity", in *Mind*: 102, 133–145.
Pawlak, Z. (1983), *Systemy informacyjne*, Warszawa: Wydawnictwa Naukowo-Techniczne.
Odrowąż-Sypniewska, J. (2000), *Zagadnienie nieostrości*, Warszawa.
Odrowąż-Sypniewska, J. (2005), "Vagueness and The Problem of The Many", in *Logic, Methodology and Philosophy of Science at Warsaw University*, edited by A. Brożek, J. Jadacki and W. Strawiński, Wydawnictwo Naukowe Semper, 117–136.
Orłowska, E. (1985), "Semantics of Vague Concepts", in *Foundations of Logic and Linguistics*, edited by G. Dorn and P. Weingartner, Salzburg: Plenum Press.
Quine, W. V. O. (1960), *Word and Object*, MIT Press.
Quine, W. V. O. (1985), "Events and Reification", in *Actions and Events*, edited by E. LePore, B. McLaughlin, Blackwell.
Sattig, T. (2013), "Vague Objects and The Problem of The Many", in *Metaphysica*: 14, 411–223.
Stojanovic, I. (2007), "Taking about Taste", in *Linguistics and Philosophy*: 30, 691–706.
Tye, M. (1996), "Fuzzy Realism and the Problem of the Many", in *Philosophical Studies*: 81, 215–225.
Unger, P. (1980), "The Problem of the Many", in *Midwest Studies in Philosophy*: 5, 411–467.
Van Inwagen, P. (1990), *Material Beings*, Cornell University Press.
Weatherson, B. (2003), "Many Many Problems", in *The Philosophical Quarterly*: 53, 481–501.
Williams, J. R. G. (2006), "An Argument for the Many", in *Proceedings of the Aristotelian Society*: 106, 411–417.
Williamson, T. (1994), *Vagueness*, Routledge.
Wójtowicz, A. (1998), "Some Philosophical Aspects of Indiscernibility", in *Incomplete Information*: *Rough Set Analysis*, edited by E. Orłowska, Physica Verlag, 381–398.

Øystein Linnebo and Stewart Shapiro

Realizability as a Kind of Truth-Making

Abstract: This paper is part of a larger project concerning potentiality in mathematics. The first and simplest case is the traditional Aristotelian notion of potential infinity. An issue much like that of truthmaking arises in our explication of one of the options for potential infinity, namely how to make sense of generalizations from that perspective. We use the traditional intuitionistic notion of realizability to resolve the issue, and to help settle the correct logic for one kind of potential infinity.

1 Background on potential infinity

Although this paper is self-contained, it is part of a larger project concerning potentiality in mathematics. The first and simplest case is the traditional Aristotelian notion of potential infinity (see Linnebo and Shapiro (forthcoming)). An issue much like that of truth-making arises in our explication of one of the options for potential infinity, namely how to make sense of generalizations from that perspective. We use the traditional intuitionistic notion of realizability to resolve the issue, and to help settle the correct logic for one kind of potential infinity.

We begin with some highlights of our account(s) of potential infinity. From Aristotle until the nineteenth century, the vast majority of major philosophers and mathematicians rejected the notion of the actual infinite. They argued that the only sensible notion is that of potential infinity – at least for scientific or, later, non-theological purposes.

In *Physics* 3.6 (206a27-29), Aristotle wrote, "For generally the infinite is as follows: there is always another and another to be taken. And the thing taken will always be finite, but always different"(2o6a27-29). As Richard Sorabji (2006), pp 322–323, puts it, for Aristotle, "infinity is an extended finitude" (see also Lear (1980, 1982)).

This orientation towards the infinite was endorsed by mainstream mathematicians as late as Gauss (1831), who in 1831 wrote: "I protest against the use of infinite magnitude as something completed, which is never permissible in mathematics. Infinity is merely a way of speaking."

A definitive change in mathematicians' orientation towards the infinite took place late in the nineteenth century, resulting in large part from pioneering work by Georg Cantor, who showed us how to make mathematical sense of completed

https://doi.org/10.1515/9783110664812-020

infinite collections or sets, and how to assign a size or *cardinal number* to such sets. Cantor's theory of infinite sets and numbers proved so elegant, insightful, and useful for mathematical purposes that it was quickly assimilated into mathematical practice, where it came to serve an important foundational role.

From this point on, the only sustained opposition to the Cantorian conception of the actual infinite came from intuitionists and constructive mathematics. They maintained that the the existence of a set – or any other kind of mathematical object – requires an explicit specification or construction. It follows that there can be no room for the actual or completed infinite. We can, with some idealization, be said to be able to construct arbitrarily large finite sets. But as finite creatures, it is out of the question that we ever complete the construction of an infinite set. It follows, intuitionists maintain, that the only permissible notion of infinite is the potential one.

We are particularly interested in some logical questions concerning potential infinity. Inspired, perhaps, by its only recent defender, many philosophers and logicians believe there is a connection between potential infinity and intuitionistic logic. Others deny the existence of any such connection. After all, thinkers from Aristotle until Gauss rejected the actual infinite in favor of the potential, but never questioned the law of excluded middle.

This current state of uncertainty and confusion concerning potential infinity is manifested in some questions raised by the acclaimed logician and philosopher William Tait (on the Foundations of Mathematics – FOM – discussion board):

> Both Hilbert and the early intuitionists have associated commitment to the actual infinite with the use of classical logic, so that, for example, the use of quantification over the integers combined with classical logic commits one to the set of integers as an actual infinity. I would like someone to explain why this is the same notion of actual infinity as Aristotle's.

We thus have the following questions:

1. If the natural numbers are merely potentially infinite, are we entitled to quantify over all of them using (at least) intuitionistic logic?

2. Does quantification over all the natural numbers with classical logic presuppose actual infinity?

Our accounts invoke a modal language. To use the heuristic of possible worlds, the idea is that, for the Aristotelian, each world is finite, but each world has "access" to another world, one with more mathematical objects.[1]

Consider, for example, the natural numbers. According to the Aristotelian, and the intuitionists, the sequence of natural numbers is merely potentially infinite. For the intuitionist, they are constructed one at a time. For both camps, we never have all of the natural numbers, existing at once so to speak. The orientation is represented as the conjunction of the following theses:

$$\Box\forall m\Diamond\exists n\,\textsc{Succ}(m,n) \tag{1}$$

$$\neg\Diamond\forall m\exists n\,\textsc{Succ}(m,n), \tag{2}$$

where $\textsc{Succ}(m,n)$ says that n is an immediate successor of m.

To be sure, the language of mathematics is strictly non-modal. Thus, when the question of the appropriate logic for reasoning about potential infinity arises, it typically does so in the ordinary, non-modal language of arithmetic.

We thus need a translation to serve as a bridge connecting the non-modal language in which mathematics is ordinarily formulated with the modal language in which our analysis of potential infinity is developed. Suppose we adopt a translation * from the non-modal language, say L, to the corresponding modal language, say $L^\Diamond$. The question of the right logic of potential infinity is the question of which entailment relations obtain in L. To determine whether $\phi_1,\ldots,\phi_n$ entail ψ, we need to apply the translation and ask whether $\phi_1^\star,\ldots,\phi_n^\star$ entail $\psi^\star$ in the relevant modal system.

This means that the right logic of potential infinity depends on two factors. First, the logic obviously depends on our modal analysis of potential infinity; in particular, on the modal logic that is used in this analysis. Included here is also the background logic for the modal language, whether it is classical or intuitionistic, or something else. Second, the logic depends on the bridge that we choose to connect the non-modal language L of ordinary mathematics with the modal language $L^\Diamond$ in which our analysis of potential infinity is given.

To summarize, the modal logic we use is S4.2, which has the S4 axioms together with

$$\Diamond\Box p\to\Box\Diamond p. \tag{G}$$

The logic is sound and complete for Kripke frames that are reflexive and transitive, and have the following "convergence" property:

1 This talk of possible worlds is, for us, only a heuristic. If we accepted the existence of possible worlds, there would be an (actual) infinity of them.

When a world w_0 accesses two other worlds w_1 and w_2, there is another world w_3 that both of the aforementioned worlds access.

$$\forall w_0 \forall w_1 \forall w_2 (w_0 \leq w_1 \wedge w_0 \leq w_2 \rightarrow \exists w_3 (w_1 \leq w_3 \wedge w_2 \leq w_3))$$

We also adopt the converse Barcan formula, as consistent with the underlying assumptions:

$$\Diamond \Box p \rightarrow \Box \Diamond p. \tag{G}$$

The translation goes as follows:

- The connectives are translated homophonically.
- An existential quantifier in mathematics ($\exists x$) is rendered as ($\Diamond \exists x$). The idea is that an ordinary mathematical statement in the form "there is a number n such that …" is to be rendered "it is possible to construct a number n such that …".
- A universal quantifier in mathematics ($\forall x$) is rendered as ($\Box \forall x$). So an ordinary statement in the form "for all numbers n, …" is to be rendered, "necessarily, all numbers (whenever constructed), …".

2 Mirroring

We now state two thereoms that relate the modal framework to its non-modal counterpart, via the translation. One is where the background modal logic is classical, and the other is where it is intuitionistic. The results support the foregoing framework.

For each formula ϕ in the non-modal language L, let $\phi^\Diamond$ be the translation of ϕ (i.e., with a $\Box$ before each universal quantifier and a $\Diamond$ before each existential quantier) into the modal language $L^\Diamond$.

Say that a formula ϕ in the modal language is *stable* if the necessitations of the universal closures of the following two conditionals hold:

$$\phi \rightarrow \Box \phi \qquad\qquad \neg \phi \rightarrow \Box \neg \phi$$

Intuitively, a formula is stable just in case it never "changes its mind", in the sense that, if the formula is true (or false) of certain objects at some world, it remains true (or false) of these objects at all accessible worlds as well. We argue elsewhere that, for the notion of potential infinity, atomic formulas should be stable (Linnebo (2013), Linnebo and Shapiro (forthcoming)). Here is our first result.

Let $\vdash$ be the relation of *classical* deducibility in a non-modal first-order language, and $\vdash^\Diamond$ be deducibility in this language corresponding by $\vdash$, S4.2, and axioms asserting the stability of all atomic predicates of the language.

Theorem 2.1 (Classical mirroring). *For any formulas $\phi_1, \ldots, \phi_n, \psi$, we have:*

$$\phi_1, \ldots, \phi_n \vdash \psi \quad \textit{iff} \quad \phi_1^{\Diamond}, \ldots, \phi_n^{\Diamond} \vdash^{\Diamond} \psi^{\Diamond}.$$

See Linnebo (2013), Theorem 5.4.

Our second result is a counterpart of this for an intuitionistic modal language. As usual, say that a formula ϕ is *decidable* in a given (intuitionistic) theory if the universal closure of $\phi \vee \neg\phi$ is deducible in that theory. Let $\vdash_{int}$ be the relation of intuitionistic deducibility in a first-order language, and let $\vdash^{\Diamond}_{int}$ be deducibility in the corresponding modal language by $\vdash_{int}$, S4.2, the stability axioms for all atomic predicates, and the decidability of all atomic formulas.

Theorem 2.2 (Intuitionistic potentialist mirroring). *For any formulas $\phi_1, \ldots, \phi_n$, ψ, we have:*

$$\phi_1, \ldots, \phi_n \vdash_{int} \psi \quad \textit{iff} \quad \phi_1^{\Diamond}, \ldots, \phi_n^{\Diamond} \vdash^{\Diamond}_{int} \psi^{\Diamond}.$$

See Linnebo and Shapiro (forthcoming), Appendix B.

These theorems have a simple moral. Suppose we are interested in logical relations between the modalized translations, in a classical (or intuitionistic) modal theory that includes S4.2 and the stability axioms (and the decidabiity of atomic formulas). Then we may delete all the modal operators and proceed by the ordinary non-modal logic. In particular, under the stated assumptions, the "modalized quantifiers" $\Box\forall$ and $\Diamond\exists$ behave logically just as ordinary quantifiers, except that they generalize across all (accessible) possible worlds rather than a single world.

We take it that this buttresses our choice of the potentialist translation as the bridge connecting actualist and potentialist theories. We also get quick answers to our two guiding questions. If the natural numbers are merely potentially infinite, we are entitled to quantify over all of them using (at least) intuitionistic logic.

More important perhaps, is an answer to the second question. Quantification over all the natural numbers with classical logic does *not* presuppose actual infinity. The first mirroring theorem shows that ordinary classical first-order logic is validated – provided it is sanctioned in the modal language.

Consider an Aristotelian notion of potential infinity. We take it this notion is based on some form of metaphysical modality, which behaves classically. Given this and the fact that Aristotle does not seem to allow any exceptions to the Law of Excluded Middle, he – and all the thinkers he inspired – are entitled to take the logic of potential infinity to be classical.

So there is no *direct* connection between potential infinity, as such, and intuitionistic logic. Motivation for the latter comes from a further orientation toward the underlying modality.

3 The modality

Actualism about the infinite unreservedly accepts actual infinities, and thus finds no use for modal notions in mathematics (or at least no use when it comes to the infinity). Actualists maintain that the non-modal language of ordinary mathematics is already fully explicit and thus deny that we need a translation into some modal language. We take it that this is the dominant view today. Actualists typically accept classical logic when reasoning about the infinite and typically also accept all of classical mathematics.

Potentialism is the orientation that stands opposed to actualism. According to this orientation, the objects with which mathematics is concerned are generated successively, and some of these generative processes cannot be completed.

Potentialists differ with respect to a *qualitative* matter. As characterized here, potentialism is the view that *the objects* with which mathematics is concerned are successively generated and that some of these generative processes cannot be completed. But what about *the truths* of mathematics? On any form of potentialism, these are modal truths concerned with certain generative processes. But how should these truths be understood?

Liberal potentialists regard the modal truths as unproblematic. In particular, there are modal truths about generative processes in their entirety, including those that cannot be completed. Consider the Goldbach conjecture. As potentialists interpret it, the conjecture says that necessarily any even natural number that is generated can be written as a sum of two primes. Liberal potentialists maintain that this modal question has an unproblematic answer – it is either true or false.

This approach to modal theorizing in mathematics is thus much like a realist approach to theorizing in general: there are objective truths about the relevant aspects of reality – modal reality in the case at hand. For our liberal potentialist, this objectivity warrants the use of some classical form of modal logic. Theorem 2.1 above, our first mirroring result, shows that, provided the modal logic is sufficiently strong, liberal potentialists are entitled to classical first-order logic in their ordinary mathematical reasoning about potential infinity.

The matter is closely related to the so-called weak Brouwerian counterexamples to excluded middle. The liberal potentialist insists that there will or will not be a sequence of twelve consecutive 5's somewhere in the decimal expansion of π. This is not, of course, because the sequence exists all at once, so to speak, as an actual infinity. Rather, the liberal potentialist notes that, in this case, the digits of the sequence are completely determined by a fixed rule, and it is this rule that guarantees that either there will or will not be the run of 5's.

Ian Rumfitt (2015), §7.4, argues against the intuitionistic use of weak counter-examples (not under that name) against classical logic. If the decimal expansion of π does contain the pattern in question, then, of course, that will be discovered eventually (idealizing, of course). Rumfitt argues that if "the pattern occurs nowhere in the expansion, *it will lie in the rule for expaning π, together with the axioms that characterize the sequence of natural numbers*, that it occurs nowhere" (p. 204). Rumfitt thus articulates and supports what we call liberal potentialism here.[2]

The *strict potentialist* goes beyond the liberal view by requiring, not only that every object be generated at some stage of a process, but also that every truth be "made true" at some stage. For example, "18 is the sum of two primes" is plausibly taken to be "made true" once the numbers up through 18 have been generated. By that stage, we have established that $5 + 13 = 18$ and $7 + 11 = 18$. There is no need to look beyond these five numbers to determine the truth of our example.

4 Generality

For the strict potentialist, an existential generalization is plausibly "made true" once a witness is generated, and shown to be a witness. That is, a sentence in the form $\exists x \Phi(x)$ is made true once a number n is generated where it is made true that $\Phi(n)$. This much does not seem particularly problematic.

By contrast, universal generalizations over the natural numbers pose a problem for strict potentialism. Again, consider Goldbach's conjecture. Since the conjecture is concerned with *all* the natural numbers – whenever they may be generated – it is hard to see how it could be "made true" at any finite stage where, after all, only finitely many numbers will have been generated.

Our goal here is to provide a modest resolution of the issue, at least for the theory of the natural numbers, using some well-known work on intuitionistic systems.

If we are serious about the merely potential character of the sequence of natural numbers, and if we are strict about the modality, then there cannot be any arithmetical truths that are "made true" only by the sequence in its entirety.

2 Since our liberal potentialist is a realist (in truth-value) about the modality, we might be tempted to call her a "modal realist", except, of course, that that label has already been taken. As noted, none of our characters is a Lewis-style realist about possible worlds, as that would require an actual infinity of worlds.

All potentialists agree that when a sequence is incompletable, there is no such thing as "the sequence in its entirety". In short, the extra demand that differentiates strict from liberal potentialism puts great pressure on universal generalizations over a potentially infinite domain. To put it baldly, it is hard to see how a generalization over all *F*'s could be "made true" at some particular stage at which most *F*'s haven't even been generated.

For strict potentialism to provide a coherent conception of potential infinity, two challenges need to be met. First, the loose talk about being "made true" needs to be made formally precise. For example, we need to explain what it is for Goldbach's conjecture to be "made true" at some finite stage, despite the conjecture's concern with numbers not generated by that stage.

Second, we need to explain how strict potentialists can make sense of universal generalizations over potentially infinite collections, such as the natural numbers. Without the ability to state and prove such generalizations, we do not have a conception of the infinite, only a draconian restriction of quantificational reasoning to finite domains.

Our aim in what follows is to show how both challenges can be met and to show that this implies that the logic of potential infinity is intuitionistic, a more nuanced answer to our second guiding question. In particular, on our account, there is no guarantee that either the conjecture or its negation will be made true by any finite stage – thus the connection to excluded middle and intuitionistic logic.

What might it be for a universal generalization over the natural numbers to be "made true" at some finite stage? Traditional intuitionism, as well as many of its contemporary defenders, provide one answer, namely to identify mathematical truth with proof, or at least with algorithmic strategies guaranteed to yield proofs.

On this strong anti-realist conception of truth, a generalization is "made true" when we produce a proof of it. And since every proof is produced at some finite stage, this satisfies the strict potentialist's requirement. The problem is the strong anti-realism on which this answer rests. The generative process is understood as a process of actual constructions, whereby mathematical objects *and truths/proofs* – which did not previously exist or obtain – are brought into being.

It is better, we think, not to rely on this anti-realism. Fortunately, another route is available, one that avoids saddling strict potentialism with the controversial anti-realist views of traditional intuitionism. We take our cue from Hermann Weyl (1921), p. 54, who writes in a discussion of whether there is a natural number that has some decidable property *P* as follows.

> Only the finding *that has actually occurred* of a determinate number with the property *P* can give a justification for the answer "Yes," and – since I cannot run a test through all numbers

> – only the insight, that it lies in the essence of number to have the property not-*P*, can give a justification for the answer "No"; Even for God no other ground for decision is available.

On this view, the truth of the universal generalization – that every number is not-*P*– has nothing to do with epistemic matters, such as proofs. Even God, who is assumed to know all the facts, cannot know facts that require running through all the natural numbers. The point is rather that, for our strict potentialist, *there are no such facts*.

Weyl's proposal, as we understand it, is that not every generalization is "made true" by the totality of its instances. When a universal generalization is true, it is instead made true by its lying *in the essence number* to have the relevant property.

By way of analogy, consider the truths that every red object is colored and that every atom of gold consists of 79 protons. These truths seem unconcerned with individual red objects or atoms of gold. They seem to be "made true" not by their instances but by *what it is to be* red or colored or gold. Likewise, we propose, there are essence-based constraints on any future generation of the objects studied by mathematics. For example, it is a constraint on the generation of natural numbers that the arithmetical successor operation be injective:

$$\textsc{Succ}(x, x') \wedge \textsc{Succ}(y, y') \wedge x' = y' \rightarrow x = y \tag{1}$$

This important arithmetical axiom is "made true" by the mentioned constraint *prior to the generation of any particular natural numbers*, or at least prior to the generation of all numbers.

Admittedly, these are deep metaphysical waters, even if not exactly those of orthodox intuitionism. In what follows, we explain the ideas in question and provide at least one precise mathematical model.

A striking feature of our explication of strict potentialism is that it leads to intuitionistic logic. This means that we provide a route from strict potentialism to intuitionistic logic that is independent of any form of anti-realism.

5 A mathematical model of truthmaking

In the case of arithmetic, at least, a good first approximation is provided by *the realizability interpretation*, going back to Stephen Cole Kleene (1945). The loose talk about what "lies in the essence of number" is to be understood in terms of (codes of) computable functions.

For natural numbers m, n, let $[m, n]$ be the code of the ordered pair of m and n. And let $\{e\}(n)$ be the result of applying Turing machine with index e to the input $\bar{n}$. More formally, let $T(e, x, y)$ be the Kleene predicate, stating that y is the

code of a complete computation of the Turing machine with index e and input $\bar{n}$. And if y is the code of a complete computation, then let $U(y)$ be its output. Then the statement that $\{e\}(n)$ is defined is

$$\exists y T(e, n, y). \tag{1}$$

And $\{e\}(n) = m$ comes to

$$\exists y(T(e, n, y) \wedge m = U(y)). \tag{2}$$

We now define what it is for a natural number e to be a *realizer* for a sentence ϕ, written $e \Vdash \phi$. The idea is that e encodes information that establishes the truth of ϕ. For present purposes, a useful metaphysical heuristic is that e functions as a "truth maker" for ϕ.

- Atomic sentences are made true by computations, and so are unproblematic. The custom is to let any natural number realize any true atomic sentence (and no natural number realizes a false atomic sentence).
- A number n realizes a sentence in the form $\phi \wedge \psi$ just in case n is the code of an ordered pair whose first element reallizes ϕ and whose second element realizes ψ.
- A number n realizes a sentence in the form $\phi \vee \psi$ just in case n is the code of an ordered pair whose first element is 0 and whose second element reallizes ϕ, or else n is the code of an ordered pair whose first element is 1 and whose second element realizes ψ.
- A number e realizes a sentence in the form $\phi \rightarrow \psi$ just in case for all numbers n, if n realizes ϕ, then $\{e\}(n)$ is defined, and realizes ψ.
- A number n realizes a sentence in the form $\neg\phi$ just in case there is no number m that realizes ϕ.
- A number n realizes a sentence in the form $(\exists x)\phi(x)$ just in case n is $[p, q]$ and p realizes $\phi(q)$.

For present purposes, the most important clause is the one for the universal quantifier:

- e realizes a sentence in the form $\forall x\, \phi(x)$ just in case for all n, $\{e\}(n)$ is defined and realizes $\phi(n)$.

That is, e realizes the universal generalization $\forall n\, \phi(n)$ just in case the Turing machine $\{e\}$ computes a realizer for the instance $\phi(n)$ when given any numeral $\bar{n}$ as input.

In terms of our metaphysical heuristic: e is a truth maker for $\forall n\, \phi(n)$ just in case e specifies a function that maps any numeral $\bar{n}$ to a truth maker for the associated instance $\phi(n)$.

Let's now regard a formula as true just in case it has a realizer or "truth maker". Since a realizer is just a natural number, this means that any true formula is made true after finitely many steps– once its realizer is generated. So our strict potentialist avoids having to wait until the end of time for at least some formulas to be made true.

Of course, there remains the question of whether this definition yields the *right* truths. For our strict potentialist, a natural measure of what is "right" is provided by the standard intuitionist theory of arithmetic, known as *Heyting arithmetic*, whose axioms as the same as those of first-order Peano-Dedekind arithmetic but where the underlying logic is intuitionistic, not classical.

Pleasingly, there is a rather straightforward theorem stating that Heyting arithmetic is sound with respect to the notion of truth that we have defined:

Theorem 5.1 (Realizability)**.** *Every theorem of Heyting arithmetic has a realizer.*

However, there are theorems of first-order *classical* Dedekind Peano arithmetic that do not have a realizer. Indeed, there are sentences that have realizers that are classical *logical falsehoods*.

In this context, Church's thesis is the following scheme:

$$\forall x \exists! y \Phi(x, y) \rightarrow \exists e \forall x \exists z (T(e, x, z) \wedge \phi(x, U(z))), \tag{3}$$

one instance for each formula Φ not containing e free. In words, an instance of Church's thesis says that if, for each x there is a unique y such that $\Phi(x, y)$, then there is a Turing machine that computes the y from the x.[3]

It is straightforward to show that every instance of Church's thesis has a realizer. But some instances of Church's thesis are provably false in classical arithmetic. Indeed, let $\Phi(x, y)$ be $[(y = 1 \wedge \exists z T(x, x, z)) \vee (y = 0 \wedge \neg \exists z T(x, x, z)))]$. It follows from excluded middle that $\forall x \exists! y \Phi(x, y)$, but, of course, no recursive function can compute this "self-halting" problem.

A. S. Troelstra (1998), §1.11, shows how to formulate realizability within the language of arithmetic, and gives a sharp result. HA* is a conservative extension of Heyting arithmetic, augmented with a term for the application of partial recursive functions, and ECT_0 is a slight extension of Church's thesis. Then HA* plus ECT_0 entails that any formula A is equivalent to the statement that A has a realizer. So

3 The connection with the more usual, informal formulation of Church's thesis is that, on the so-alled BHK reading of the intuitionistic quantifiers, a formula in the form of the antecedent, $\forall x \exists! y \Phi(x, y)$, means something like "given any x, one can effectively find a unique y such that $\Phi(x, y)$". And the consequent is that there is a Turing machine that computes this value.

all we are entitled to assume concerning truth making at a world is intuitionistic logic (augmented with Church's thesis).

Let us now kick away the ladder of this talk about worlds in favor of just the modal operators, representing the generative modality and based on S4.2. It follows that this logic must be based on intuitionistic logic, not classical. Thus, for the strict potentialist, the appropriate modal system is one based on intuitionistic S4.2.

As before, the next step is to apply the potentialist translation to provide a bridge back to the non-modal language of ordinary mathematics. The result of doing so has already been established by Theorem 2.2, on intuitionistic potentialist mirroring. As we recall, the theorem ensures the following:

$$\phi_1, \ldots, \phi_n \vdash_{int} \psi \quad \text{iff} \quad \phi_1^{\Diamond}, \ldots, \phi_n^{\Diamond} \vdash_{int}^{\Diamond} \psi^{\Diamond}.$$

This confirms a claim by Dummett and others, which has so far never been properly substantiated, namely that only intuitionistic quantification, not classical, is permitted over a domain that is potentially infinite (or "indefinitely extensible") – provided the potentiality is strict.

6 Limitations, and future work

These results, as pleasing as they are, do not readily extend to other, richer mathematical domains. For one thing, recall that Theorem 2.2, on intuitionistic mirroring, requires that atomic formulas are all decidable. This is an easy theorem for arithmetic, but it fails with the real numbers (where identity is not decidable).

Second, our account of truthmaking, via realizability, requires that each object in the domain – each natural number in the case at hand – has a canonical name. This, too, fails once we go beyond the natural numbers.

And third, our account uses recursive function theory, a rich and well developed account of computation over natural numbers and other finitary things like strings. There is interesting body of work on computations involving objects richer than natural numbers or strings. And there is an interesting body of work on realizabilty for richer intuitionistic theories, such as real analysis and set theory. But it is not at all clear, at this stage, whether this work can be used to provide an account of generality in those contexts.

Bibliography

Aristotle (1941), *The Basic Works of Aristotle*, edited by R. McKeon, Random House.

Cantor, G. (1883), *Grundlagen einer allgemeinen Mannigfaltigkeitslehre. Ein mathematisch-philosophischer Versuch in der Lehre des Unendlichen*, Leipzig: Teubner.

Cantor, G. (1887), "Mitteilugen zur Lehre vom Transfiniten 1, II", in *Zeitschrift für Philosophie und philosophische Kritik*: 91, 81–125, 252–270; and 92, 250–265. Reprinted in Cantor (1932), 378–439.

Cantor, G. (1932), *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts*, edited by E. Zermelo, Berlin, Springer.

Gauss, K. F. (1831), "Briefwechsel mit Schumacher", in *Werke*: Band 8, 216.

Kleene, S. C. (1945), "On the Interpretation of Intuitionistic Number Theory", in *Journal of Symbolic Logic*: 10, 109–124.

Lear, J. (1980), "Aristotelian Infinity", in *Proceedings of the Aristotelian Society*: 80, 187–210.

Lear, J. (1982), "Aristotle's Philosophy of Mathematics", in *Philosophical Review*: 41, 161–192.

Linnebo, Ø. (2013), "The Potential Hierarchy of Sets", in *Review of Symbolic Logic*: 6, 205–228.

Linnebo, Ø. and Shapiro, S. (forthcoming), "Actual and Potential Infinity", in *Noûs*.

Rumfitt, I. (2015), *The Boundary Stones of Thought*: *An Essay in the Philosophy of Logic*, Oxford, Oxford University Press.

Sorabji, R. (2006), *Time, Creation and the Continuum: Theories in Antiquity and the Early Middle Ages*, Chicago, University of Chicago Press.

Troelstra, A. S. (1998), "Realizability", in *Handbook of Proof Theory*: *Studies in Logic and the Foundations of Mathematics*, edited by S. R. Buss,, Amsterdam: North Holland Publishing Company, 407–473.

Weyl, H. (1921), "Über die neue Grundlagenkrise der Mathematik", in *Mathematische Zeitschrift*: 10, 39–79.

Part VI: God, Theodicy, and the Best World

Nicholas Rescher

Optimalism in Explaining the Nature of Things

Abstract: In the final analysis, a cogent explanation of the world's existence in the standard factual order of experiencing some facts by means of others is infeasible. For here the premisses are themselves bound to be part of the problem. At this stage of natural deliberations there will have to be a shift from facts to values, via a Leibnizian explanation that pivots existence on value and thereby grounds the world's existence and nature in this being for the best.

1 The Leibniz question

Among the fundamental issues of metaphysics is the ultimate question put on the agenda of the field by G. W. Leibniz: "Why is there anything at all?"

Before this question can be meaningfully addressed, some essential clarification must be provided.

To begin with, what sort of "thing" is to be at issue in this question? Are *numbers* to count as "things"? Then reasons of necessity will do the job. Or again, if *facts* (states of affairs) are to count as "things" then the answer is once more straightforward: there are such things because their necessity has it so. Some things – number and facts – necessarily exist. And there is also – according to many thinkers – yet another necessary existent, viz. God. And so as long as "things" like facts, and numbers (let alone deities) are allowed into the range of relevancy, the answer to the Leibnizian question is simply: "Because it has to be so and cannot possibly be otherwise."

However, this sort of consideration is really irrelevant. For the pivot of Leibnizian concern is actually

– Why is there something *contingent* – something whose existence is not necessary?

Thus in the context of present concern, the necessary being of abstractions is beside the point and what is really at issue there is the matter of *contingent existence*. At bottom, that initial question is intended to ask: "Why is there a realm of contingent existence – a real world with concrete objects in it? Why are there actually spatio-temporal reals when there might possibly not be?"

And of course once this issue is resolved satisfactorily there arises the further and no less problematic question:

https://doi.org/10.1515/9783110664812-021

– Why is it that *this* contingent order of things exists rather than some possible alternative?

2 Distributive explanation cannot do the job?

To begin with, it should be clear that an ultimate theory of explanation – one that addresses contingent existence-at-large must be holistic: it must address the integrity of a collective whole, the world. To be sure, some theorists endorse what has come to be called the "Hume-Edwards Thesis" that: *If the existence of every member of a set is explained, then the existence of the set is thereby explained*.[1] And they then propose to resolve the Leibnizian question seriatim, by explaining the existence of every existent through a causal explanation of its origination.

However, the fallacy at issue is not too difficult to see. For consider the following two claims:

– If the existence of every sentence of a paragraph is explained, the existence of that paragraph is thereby explained.

– If the existence of each note of a symphony is explained, the existence of that symphony is thereby explained.

Both of these claims are clearly false as they stand. On the other hand, contrast these two with the following cognate revisions:

– If the existence of every sentence of a paragraph *as a sentence of that particular paragraph* is explained, then the existence of that paragraph is thereby explained.

– If the existence of every note of a symphony *as a part of that particular composition* is explained, then the existence of that symphony is thereby explained.

Both these theses are indeed true – but only subject to that added qualification. After all, to explain the existence of the spouses is not automatically to achieve an explanation of the marital couple, seeing that this would call not just for explaining these participants distributively but their collectively coordinated co-presence

1 Rowe (1970), p. 153. On this principle in its relation to the cosmological aspect for the existence of God see Rowe (1975). See also Gale (1991) and Pruss (1988).

in the structure at issue. And the case is just the same with the Hume-Edwards thesis.

Explanatory invocation of the Hume-Edwards Principle fails to heed certain critical *conceptual* distinctions that are readily brought to light by means of a bit of symbolic machinery. So let us adopt the following abbreviations:

- p @ q for "p [is true and] provides an adequate explanatory account for q", where the variables p and q range over factual claims.
- E!x for "x exists", where the variable x ranges over possible existents.

Since the variable x ranges over existents, we have it that $(\forall x)$E!x.

On this basis it is readily brought to view that the form of the statement "Everything has an explanation" or "There is an explanation for everything" admits of two very different constructions:

(1) $(\forall x)(\exists p)(p$@E!$x)$
Distributive explanation: "There is some case-specific explanation to account for each and any individual existent."

(2) $(\exists p)(\forall x)(p$@E!$x)$
Collective explanation: "There is one single comprehensive explanation that accounts for all existents – the entire totality of them."[2]

It is clear that very different questions are at issue and very different matters at stake with distributive and collective explanations. For distributive explanations explain the fact *that* every member of a certain set has the feature *F*; collective explanations account for *why* it is that this is so. And explaining how it is that all members of the club are male – which could be so by fortuitous circumstances – does not accomplish the job of explaining *why* this is so (e.g. because the bylaws require it). In posing different questions we must be prepared for the possibility of different answers.

Specifically a significant distinction is at issue here. Why-is-there questions are apt to be equivocal. The interrogative "Why are there lions – why is it that lions exist in nature" can ask two very distinct questions:

1. How has it come about that there are lions – that lions have come to realization in the evolutionary scheme of things?
2. Why is it that the process of organic evolution are such as to make for the emergence of lions?

2 Note that neither of these is the same as $(\exists p)(p$@$(\forall x)$E!$x)$ which obtains trivially given the symbolic conventions adopted here.

Accordingly, a why-is-there question can take either a local and proximate form or a global and ultimate one. "Why is there a moon" can ask for an account in the context of the development of the solar system, or it can ask the cosmic and virtually metaphysical questions: Why is cosmic evolution so structured that this system and its moon has come about. The former, proximate questions asks for *process*; the latter ultimate questions asks for a *reason*.

Distributive questions address the issue locally and can be resolved by proximate answers; ultimate questions ask for more and have to be addressed globally. Accordingly, the Hume-Edwards Thesis will be of no real avail in our explanatory quest. One has to look elsewhere.

3 The need for oddity: Abandoning causality

One key point was made by Leibniz long ago:

> The reasons for the entire world [must] therefore lie in something extramundane, different from the chain of states or series of things whose aggregate constitutes the world. … So [to account for the world's being] there must exists something which is distinct from the plurality of beings, or from the world.[3]

In explaining the being and the nature of actual concrete existence-as-a-whole we cannot have explanatory recourse to any aspect of the being and nature of reality itself. To do so would be to "beg the question" – to make use in giving an explanation of some part, feature, or aspect of the very thing that is to be explained. And of course this mode of explanation cannot function effectively in the present context. For any causal explanation carries us back to the starting point: the presupposition of this or that existent. But the questions at issue puts this very circumstance into question. One cannot coherently invoke the existence of *something* in trying to explain the existence of *anything* whatsoever. In explaining the internality of the whole of real existence one must go outside this realm.

It would accordingly be absurd to ask for some sort of *causal* account for reality-as-a-whole. Causality, after all, is a world-internal process: its functions show how some world-integral things and conditions arise out of others. It is the sort of account we use to explain how acorns yield trees and how lion parents produce baby lions. Causality is a matter of ultra-world agency and requires world-internal inputs to do its work. It is not the sort of resources that would possibly be

3 Leibniz (1969), p. 487.

called to account for the world itself and to explain the origination of the totality of existents.

The pivotal lesson of the preceding deliberations is that one cannot adequately explain contingent existence-at-large by an appeal to the nature of existence itself. In the final analysis the nature of contingent existence must be explained not on the basis of existing things or substances, but rather in the operation of principles that function with respect to the manifold of possibility.

Its formulation at this level of synoptic generality marks the why-this-world? as decidedly nonstandard. For standard questions of existence-explanation proceed in causally putative terms. The reason that X exists is that there exist other items Y_1, Y_2, ..., Y_a which interact causally so as to engender X. In standard existence explanations, what exists emerges through the causally productive machinations of other existents. But this sort of thing clearly will not do in the present context.

The question of existence-in-general cannot be dealt with as one of the standard generative sort that asks for the existence of one thing to be explained causally in terms of the existence and functioning of another. We cannot say "Well there's X in the world, and X explains the existence of things" because this simply shifts the issue to X, which after all is itself an existent. If we want *global* explanations of existence of things in the world, we are going to have difficulty in getting them from existential premisses pertaining to what the world is like. Does this mean we cannot get them at all?

And so, with ultimate questions about existence-at-large eccentricity is unavoidable. Those ultimate, totalistic and holistic questions are altogether extraordinary. Usually when we ask about things and their conditions we are after a developmental account – how they got to be so given a process of transformation from some earlier condition. This standard sort of issue-resolution is clearly impossible in the present case. The fact of it is that when we ask an extraordinary question we must be prepared for an extraordinary and indeed seemingly bizarre answer. For if an altogether basic condition of things is to be explained this cannot be achieved appropriately on the basis of the machinations within the realm of existing things.

4 A two-fold turning

To secure our explanatory basis for contingent existence at large a radical step becomes inevitable. For here one has to redirect one's line of thought in two directions, from actuality to possibility and from fact to value. Let us consider how

these reorientations are to work. To begin with there is the turn to metaphysical possibility.

To account for the being of contingent-existence at large one has to put the burden explanation on something that is itself entirely outside the realm of contingent existence. With such an "ultimate" question the explanatory appeal has to move outside the entire realm of existential fact.

But where can we possibly go if this condition is to be met?

But where can one possibly look for explanatory resources of the realm of actuality, if "what there is", is not available?" The answer is clear: we must look to the realm of possibility, of what *can possibly be*. And to have any explanatory traction here, we must invoke the conception of value – of what there ought to be. Thus to resolve the problem on the metaphysics of existence we must ultimately turn to a metaphysics of value.

The existential non-variety of the entire domain as a whole cannot be cogently explained by invoking some feature of its existential content. If there is to be an acceptable explanation its probative basis must lie wholly outside this domain. It cannot be done within the realm of *things* or *substances* at all, but must step outside to proceed on the basis of some sort of *principle*.

When we are to explain some actual condition of things without involving any other actual conditions of things, we are obviously facing a very tall order. And our room for measure is extremely limited. For if we cannot explain actualities at large in terms of actualities, we have little alternative but to explain them in terms of possibilities. What is thus called for here is a principle of explanation that can effect a transit from possibility to actuality, and thereby violates the medieval precept *de posse ad esse non valet consequentia*.

Then too there is the further turn to eliminative valuation. It arises from the problem: If an adequate explanation of contingent existence is achievable only in terms of reference lying outside the realm of necessity and also outside the realm of concrete existence and contingent fact, then where can it possibly go? And the only conceivable answer here is this: it must go entirely outside the realm of fact to that of value.

To achieve a synoptically ultimate explanation of the domain of contingent (existence/reality) we thus have to shift to another domain of deliberation altogether – and move outside of the evidential *realm of what* is to the normative *realm of ought to be*, from actuality to value.

And to realize this transition we must shift from the sphere of production to that of elimination. We must effect a revolutionary shift in the orientation of thought from productivity to reducibility, from fact to value, from actuality to possibility.

In sum it must implement the idea that contingent reality is what it is because that is somehow for the best. It must, that is to say, explain existence in terms of value and take what might be called the axiological turn. And once again, the key point was made by Leibniz long ago:

> Even if the world is not necessary [absolutely or] metaphysically, in the sense that its contrary would imply a contradiction or logical absurdity, it is nonetheless necessary physically [or evaluatively], determined in such a way that its contrary would imply imperfection or moral absurdity. And thus as possibility is the principle of essence, so perfection or degree or essence is the principle of existence.[4]

Granted, this sort of thing may sound strange if not bizarre. But the reality of it is that in asking for an explanation of contingent-existence-as-a-whole one is posing a decidedly extra-ordinary question, and when you insist upon doing this, you must be prepared to a decidedly extra-ordinary answer. In this context noting the bizarre nature of the answer is not an objection to it but the acknowledgment of a *necessary* condition for its adequacy.

5 The crux is not causal production but possibility elimination

One must reckon with the situation that an ultimate account of reality-as-a-whole has to proceed not in terms of causal production but in terms of possibility-elimination based on evaluative considerations. Let us examine how this approach would have it work.

The crux of the reasoning required here lies in the Sherlock Holmes Principle: "When you have eliminated all untenable possibilities, whatever remains, whatever its features or other aspects, will be real and actual."[5] Here we move from elimination to actualization, from disqualification from the realm of possibility to entry into the domain of the real.

Now in the matter before us, this possibility-elimination cannot proceed *causally*. It has to proceed *normatively*. Those eliminated possibilities are out not because some creative agent an agency chooses to eliminate them, but because they are inherently unworthy – outranked and outflanked by far superior alternatives.

4 Leibniz (1969), p. 488.

5 Doyle (1892).

Natural explanation functions with regard to nature – to existence-in-the-world. But metaphysical explanation at the level of things-in-general have to function at the level of possibility. And here we come to the fundamental law of metaphysics: *Inferior alternatives are ipso facto unavailable for realization*. Inferior merit is existentially disqualifying. And this principles carries a crucial corollary: *Reality is optimific*. And so the answer to the question of what explains the elimination of the inferior alternatives lies in a metaphysical Optimality Principle: *Given an exhaustive range of possible alternatives, it is the best of them that is actualized*. And on this basis the answer to our initial question "Why is there something contingent – something whose existence is not necessary" is simply: "Because it is for the best that this should be so."

So what we have here is the Leibnizian Turn of shifting the explanatory strategy at work from the order of descriptive fact to that of normative evaluation.

6 Explaining the optimality principle: Self-explanation as the privot

But what is it that could account for a Leibnizian principle of optimality? What sorts of considerations could possibly provide for the justifactory validation of optimalism? Why should it be that such a principle obtains? Why should what is for the best be actual? The answer here lies in the principle itself. The principle is literally self-explaining. Realization of the Optimality Principle is itself the best alternative in accounting for the prevailing order of things.

What is at issue here is something of a principle of the coordination of explanatory and ontological merit: the best candidates for explanation and the best candidate for actualization are seen as being one and the same: the alternative that is the best for endorsement is that which is best for realization.

But is this reasoning not invalidated through circularity?

By no means! For at this stage circularity is not vicious but virtuous: it is not a flaw but an essential asset. The fact is that any *ultimate* explanation must be self-sustaining: it must rest on a principle that is self-validating. For if the validity of the principle rested on something else – some deeper and different rational of validation – then by hypothesis it would not be ultimate but would through this very circumstance be flawed.

And of course the optimality principle indeed has this feature of self-support. For the obviously appropriate answer to the question "But why is it that such a principle of optimality obtains" is simply "The principle obtains because that is for the best". Clearly the principle is self-sustaining and self-explanatory. And in

this present cast this is not vitiating circularity but an essential aspect of the problem – decidedly virtuous circularity.

The circumstance that an ultimate principle must be self-sustaining reinforces acknowledgment of the necessary shift from descriptive facts to normative values. For matters of descriptive fact cannot function self-sustainingly without vitiating circularly. But there is no sound reason for holding that value considerations cannot do so.

The Principle of Optimality has a *raison d' être* alright. But it lies in its own nature. For it is, in the final analysis, for the best that the Principle of Optimality should obtain. After all, there is no decisive reason why that explanation has to be "deeper and different" – that is, no decisive reason why the prospect of *self-explanation* has to be excluded at this fundamental level.[6] After all, we cannot go on putting the explanatory elephant on the back of the tortoise on the back of the alligator ad infinitum: as Aristotle already saw, the explanatory regress has to stop somewhere at the "final" theory – one that is literally "self-explanatory". And what better candidate could there be than the Optimality Principle itself with the result that the divisions between real and merely theoretical possibilities is as it is (i.e., value based) because that itself is for the best?[7]

That Law of Optimality to the effect that "whatever possibility is for the best is ipso facto the possibility that is actualized" is certainly not a logico-conceptually necessary truth. From the angle of theoretical logic it has to be seen as a contingent fact – albeit one not about nature as such, but rather one about the manifold of *real* possibility that underlies it. Insofar as necessary at all, it obtains as a matter of ontological rather than logico-conceptual necessity, while the realm of possibility as a whole is presumably constituted by considerations of logico-metaphysical necessity alone.[8]

6 After all, there is no reason of logico-theoretical principle why propositions cannot be self-certifying. Nothing vicious need be involved in self-substantiation. Think of "Some statements are true" or "This statement stakes a particular rather than universal claim".

7 Optimalism is closely related to optimism. The optimist holds that "Whatever exists is for the best", the optimalist maintains the converse that "Whatever is for the best exists". However, when we are dealing with exclusive and exhaustive alternatives the two theses come to the same thing. If one of the alternatives A, A_1,..., A_n must be the case, then if what is realized is for the best it follows automatically that the best is realized (and conversely).

8 The operative perspective envisions a threefold order of necessity/possibility: the logico-conceptual, the ontological or proto-physical, and the physical. It accordingly resists the positivistic tendency of the times to dismiss or ignore that second, intermediate order of considerations. And this is only to be expected since people nowadays tend to see this intermediate realm as predicated in value considerations, a theme that is anathema to present-day scientism.

In the end, we must expect that any ultimate principle must explain itself and cannot, in the very nature of things, admit of an external explanation in terms of something altogether different. The impetus to realization inherent in authentic value lies in the very nature of value itself. A rational person would not favor the inferior alternative; and a rational reality cannot do so either.

So what has to be at work here is a proto-ontological law to the effect that under such-and-such conditions various theoretically available possibilities become eliminated (i.e., realization-ineligible) as real possibilities by virtue of evaluative inferiority. And such a process will have to continue to operate in the possibilistic domain until at last only one privileged alternative remains. What we have here is literally a struggle for the survival of the fittest, but now with matters being fought out not among competing actuals but among competing possibilities – that is between theoretically available possibilities and metaphysically acceptable ones.

7 The standard of metaphysical value: Noophelia and the pivotal role of intelligence

Optimalism's pivotal idea is that value explains reality – that the best available alternative is going to be actualized – will spin like a useless gear that fails to engage the machinery of explanation unless and until the standard of evaluative that is at issue is identified. Only then will this "axiogenesis" approach acquire any explanatory traction. And so, at this point the sixty-four dollar question becomes: What sort of considerations can serve as the determinant of existential fitness here? What renders one world-arrangement superior and existentially more qualified than another? Until this issue is resolved, there is no way to comprehend the principle of optimality.

The problem of ontological merit becomes unavoidable at this point. And it is clear that one cannot just optimize, any more than one can just maximize or minimize. For one has to optimize *something*, some feature or aspect of things. But if this merit-indicating factor is to be something that is self-validating and self-sustaining then the clearly most promising candidate would seem to be intelligence itself – that is to say the overall status and standing of intelligent beings at large. Any rational being is bound to see the loss of reason as a supreme tragedy. For an intelligent being – a rational creature – intelligence itself must have a prime place high on the scale of values. It, accordingly, is intelligence and rationality as such that best qualifies as the self-sufficient standard of value that will have to be

at issue "for the best" will have to be construed in terms of what is best for the enhancement and diffusion of intelligence in the cosmos.

Accordingly, the optimalism envisioned here is oriented at optimizing the conditions of existence for intelligent beings at large. And at the cosmological level such an optimalism militates towards a universe which

- provides for the chance and randomness through which alone intelligent beings can emerge in the world through evolutionary processes based on chance-conditioned variation and selection.
- provides for the chance-conditional novelty and innovation needed to provide an environment of sufficient complexity to be of interest for intelligent beings.
- provides for the order of regularity and lawfulness needed for a universe sufficiently orderly and to allow complex creatures to develop and thrive.
- provides for a lawful order in the modus operandi of nature sufficiently simple to be understood by imperfectly intelligent beings as a basis for grounding their decisions and actions in a complex world.

The arrangements of an intelligently contrived universe must, in sum, manage things in a way that rational creatures would see as optimal from the vantage point of their own best interests as rational beings. And so an optimal world, in the cosmological sense presently at issue, is one that achieves a condition of optimalization under constraints – these constraints being a manifold of natural law favorable to the best interests of intelligence – that is, intelligent beings at large in the overall scheme of things.

And so optimalism is a theory of rational systematization that grounds the explanation of the world's facts through a process of optimization subject to constraints – the constraints being the projection of a lawful order of things in which intelligent beings can emerge through evolutionary processes of a sort that affords them the opportunity to thrive physically and to progress cognitively. In effect we have: *optimalism* + *noophelia* = *axiogenesis*. The theory accordingly explains concrete reality's existence and nature in terms of their providing for a natural domain that is conducive to the success of intelligent beings in the world (which is not quite to say that such success is guaranteed).

But how can one say that reality is user-friendly for intelligent beings given the amount of human suffering in this world? This crucial question poses the so-called Problem of Evil and calls for another range of deliberations altogether. It requires separate treatment elsewhere.[9] Only one brief reminder will have to do here, viz. that in world improvement we face the prospect that in the world system

9 The problem is addressed by Leibniz in his *Theodicy*. It is also treated at length in Rescher

as with a bio-system in matters of medicinal interactions improving matters at point *A* may well call forth even greater collateral damage at point *B*.

8 A digression on explanatory homogeneity

The presently contemplated axiological explanation is (or at any rate seems to be) inconsistent with what has been called "The Principle of the Externality of Explanation" to the following effect:

> It is a necessary truth that: If it is contingently true that items of a certain kind *K* exist, then any [adequate] explanation of this fact (viz. that such items exist) must appeal to or involve items that are not of the kind *F*.[10]

This principle seems suitable in relation to many physical/material kinds. Thus if one had to explain the presence of mice in the world in kind-homogenous terms, with parent mice always needed for mice offspring, this would require an eternal and uncreated Aristotlean universe, seeing that the production of mice takes time! So – clearly – mice have to emerge from some other kind of thing. But mice are not typical here, and this sort of thing is not going to hold for contingent existents at large. For viewed more generally as regards kinds of beings this "principle" is decidedly problematic because counter-examples are readily available. Thus let "kind *F*" be defined as:

– (written) English words not containing the letter *E*.

The existence of such words is presumably contingent. Yet it is obviously false to say that an adequate explanation of the existence of such English words "must appeal to or include" words that are not *E*-containing, and cannot be provided by means of *E*-containing words alone.

Or let "beings of kind *F*" be material, physical things and let it be that physical materiality began with a "big bang" whose occurrence at the start of an *open interval* of time was the beginning of a physical universe. Then the existence of physical material throughout cosmic history, albeit contingent, can be explained adequately in terms of the materiality of earlier states – ad infinitum.

(2011).

10 Compare Van Inwagen (2015), p. 40.

Again let the "beings of kind *F*" be contingent beings at large which of course there will (contingently) be. To say that as a matter of principle explaining their existence requires an appeal to non-contingent (i.e., necessary) beings certainly begs some very large philosophical questions.

But the general point should be clear enough. To explain the existence of words (at large) we need not exist from the realm of words. In explaining the existence of ideas or concepts we can remain within the realm at issue. To explain facts about nature we need not necessarily go outside this domain.

What the principle denies is the prospect of an *homogeneous* explanation of contingent existents via items of their own kind. But this is hardly a matter of absolute necessity. In a universe that would, like Aristotle's, be eternal, uncreated, and ontologically stable, the ongoing presence of stars in the cosmos could adequately accounted for by the fact that they have always been there. What further explanation would or could – in the circumstances – be required?

Given such considerations, a conflict with "The Principle of the Externality of Explanation" should not be seen as an impediment to an axiological explanation of the world.

9 Circularity issues

In its doctrinal validation, optimalism is self-reliant, cyclic, even circular. But this is not necessarily vitiating. Circular reasoning is not by nature vicious.

On analogy with the moral dictum "Honesty is the best policy" one might offer the epistemic dictum: "Veracity is the best policy." Specifically one might endorse the following principle:

(P) If *p* is the best cognitive option – the best thing to maintain in the prevailing circumstances – then *p* can and should be accepted as true.

But is (P) itself correct and true? A case for saying so can be made out as follows:

(1) Maintaining (P) is itself our best option in the prevailing circumstances. (We cannot reasonably expect to do better than our best.)

(2) Given (1) it follows from (P) that (P) is to be accepted as true.

Note that here (P) is itself invoked on the course of its own validation, so that the proposed (P)-validation is in fact circular. Does this invalidate it?

The employment of a contention in the course of its own validation widely rejected as a matter of "begging the question". Nevertheless, the idea that such a

process of "circular" reasoning is thereby inappropriate and vitiating – et alone *vicious* – is very questionable.

Cyclic reciprocity is the lifeblood of epistemology. We assess the meaning of words in the light of the truths they are used to convey and endorse truth on the basis of what their words mean.

Circularity of this sort is not – or should not be seen – as automatically inappropriate. In many cognitive contents validation is not indirectional from for secure premisses to inferred conclusions but is determined on the light of feedback processes of cyclic and reciprocal harmonization and systematization. We assess the reliability of reports on the basis of the trustworthiness of their species and assess this trustworthiness in the light of the circumstances of these reports.

Consider the thesis:

> In validating a statement *S* it can sometimes be appropriate to make use of that very statement *S* itself as a premiss of the reasoning.

The question before us is whether this thesis is true. And in exploring this issue we shall proceed not by considerations of general principles, but rather by way of illustrations and examples.

It is not all that difficult to provide instances of self-certifying theses – claims that are cogently self-substantiating. Thus consider

- There are truths.
- This is an English sentence.
- Some English sentence begin with "some".
- Some true statements are about statements.
- Meaningful statements can mention science.
- Negative claims are not invariably false.
- Truths can be stated briefly.

All of these claims are true and are, in fact, self-certifying. The reasoning at issue here unfolds as follows:

- The claim that there are or can be items of a certain kind can be validated by adducing an item of this kind as a validating instance.
- Thesis *T* claims that there are items of kind *K*.
- Thesis *T* is itself an item of kind *K*.
- Theses *T* can be validated by adducing itself as a validating instance.

Such a line of reason pivots the establishment of a claim by providing that claim itself as a demonstrating fact. It is thus clearly circular but is equally cogent and unproblematic.

As these deliberations indicate, self-substantiating claims cam serve to show that the circularity of self-support need not be inappropriate (let alone invalidate) a claim but can be the raison d'etre a proposition's truth. For clearly, these claims about themselves constitute conclusive evidence for their own truth (and need no external support). There is circularity alright but it serves to validate rather than disrupt.

Consider the following situation:

List I	List II
– Exactly two statements of List II are true.	– Exactly two statements of List I are true.
– 1 + 1 = 2	– 1 + 3 = 3
– 2 + 2 = 5	– 2 + 3 = 4

Note here that the first statement of List I constrains the truth of statement one of List II which in turn constrains the truth of that former statement. In this situation that statement figures crucially in establishing its own truth. After all, suppose that it were not true, then List II will have to contain 0 or 1 or 3 true statements. But 0 and 3 are out, given what statements two and three of List II affirm. Hence statement one of List II must be true. And this constrains the truth of statement one of List I.

What we have here is yet another example of the appropriateness of employing as statement itself in an order to establish its truth.[11]

And what holds for validation holds for explanation to as well. Consider the explanatory question: "Why did Henry spell CAT as KAT in his letter to us of yesterday?" It is as cogent an explanation as we could possibly ask for that:

– Henry *always* spells CAT as KAT.

But this explanation is clearly circular: the very fact to be explained is incorporated in the explanatory account we are being offered.

As such examples show, propositional substantiation need not be linear and other-dependent but can inhere in a proposition's own substantiative context. In

11 Something of an anomaly is afforded by the claim: "There are falsehoods." For unlike "There are truths" this contention is not self-exemplifying. But it is nevertheless self-certifying. For what it says cannot possibly be false, since if it were, this statement would, ipso facto, be validating itself. In establishing the truth of this claim we do not accept but rather deny it.

suitable cases probative involvement can be rationally cogent and appropriate. Circularity need not be vicious: there can be virtuous as well as vicious circles.

Circularity is vitiating only within certain purposive contexts. For if the task at hand is to establish the credibility (acceptability) of *p* as a conclusion in reasoning, then it would be inappropriate to presume its availability among the premisses. Clearly if I want to persuade you of something I cannot properly proceed on the presumption that you *already* concede it to me. But this dialectal defect is certainly no logical flaw. For deductive logic has no quarrel with circularity: it has no hesitation about evolving the validation of the premiss *p*-and-*q* to the conclusion *p*, or from *p* to *p*-or-*q*.

It becomes important in this context to distinguish between demonstrative substantiation and persuasion. If I want to establish something as a matter of true fact, then its available as a factor in reasoning must sometimes be conceded. To investigate the truth-status of a claim *p*, I may have to presume its availability (as least for "the sake of discussion") as a basis for examining what follows. Thus consider a self-referential statement on the order of:

(1) This very statement (S)-itself, makes a meaningful and intelligible assertion.

If I were not prepared to presume the truth of (S) from the very start, there would be no point in continuing the inquiry.

And this situation obtains in many other contexts of circular reasoning as well. Thus since object of a definition is to eluciate the meaning of an expression in terms of others whose meaning is presupposed as known, we cannot resort to this to-be-defined expression in the course of the definition, since in so doing we defeat the prospect of realizing this object. Or again, the object of a recipe is to explain how to make something whose production is not otherwise understood, and by assuming the prior availability of this very thing we preclude the realization of this object.

Vitiating circularity occurs whenever the processual purpose at issue is to produce something that is also needed as an input for its own operations. But here the flaw is practical and pragmatic. For throughout these cases we are dealing with a process aimed and at realizing a certain objective whose attainment is defeated by circularity.

10 Reason's self-reliance is not vicious but virtuous

The only satisfactory explanation for anything – even for the existence of intelligence and its requirements – will have to be an intelligent explanation. In taking intelligence to provide its own ultimate explanatory basis we proceed in a way that is cyclical and indeed even "circular". But this simply reflects the structural coherence of rational systematization. And there is nothing viciously self-defeating about such self-reliance. For while vicious circularity stultifies by "begging the question", virtuous circularity merely coordinates related elements in their mutual interlinkage. The former pre-assumes what is to be proved, the latter simply shows how things are connected together in a well-coordinated and mutually supportive interrelationship.

And this is crucial kin the present range of deliberations. For to be able to afford an adequate resolution of our ultimate question the principle at work cannot rest on further extraneous considerations. For the question of why the truth of things is what it actually is will arise with respect to the principle itself, and if it is to resolve such matters it must do so with respect to itself as well. It must, in short, be self-sustaining and self-grounding. Otherwise the requisite ultimacy will thus be achieved.

The only validation of rational intelligence that can reasonably be asked for – and the only one worth having – must lie in considerations of the systemic self-sufficiency of reason – its endorsement on the basis of rational considerations. And such self-endorsement is not problematic but altogether appropriate.

After all, even to ask the question "Why should it be that reality is intelligible?" is already to manifest one's commitment to the principle at issue, since asking this question is to expect an answer – and a cogently intelligent one at that.

There is simply no satisfactory alternative to using intelligence in its own explanation. For when a self-validating principle of explanation is needed, then intelligence and reason appear on the scene as ready volunteers. Noophelia – intelligence favoring – accordingly provides a natural pivot for the presently envisioned optimalism.

The explanatory chain at issue here runs from the productively causal laws of nature to the eliminatively normative laws of metaphysics – from the actualistic to the possibilistic dimension and from the descriptively factualistic to the normatively axiological. A radical shift of perspective is at issue. Accustomed as we are to living and thinking within the sphere of actuality, such a change of perspectival context is bound to seem extra-ordinary. It is bound to seem strange by ordinary

standards, because those ordinary standards just do not and cannot apply at this level of deliberation. Here are face a question whose inherent nature is too far out of the usual range existence explanation that only an answer that lies entirely outside the box of accustomed thinking can even begin to address the matter in a meaningful way.

11 The perspective of theology

Questions of priority can be puzzling. Which comes first, the chicken or the egg? Or, at a more profound level, God or axiology? Is reality optimific because a benign God chooses to institute the Principle of Optimality to the effect that what is for the best will be actual? Or does God exist because it is for the best that it should be so and the Optimality Principle accordingly renders his existence actual?

The most promising resolution to this priority puzzle is to dismiss the any and all commitment to priority and instead insist on rational coordination. The situation of cause and effect is analogous: there can be no effect without a cause nor any cause without an effect – the two conceptions are coordinate and conjoint without logical priority one way or the other. And this is the case in my present circumstances as well.

Since *God is love* – a love that includes prizing, cherishing, valuing – it transpires that with him the salient aspect of value and being are inseparable: divinity and optimality are "joined at the hip" so to speak. The self-sufficiently of optimality and the self-engenderment of a *causa sui* divinity could be seen simply different aspects of one selfsame fundamental factor: the fusion of optimalism with the idea of a loving God.

Noophelia is not only compatible with but actually congenial to theism. After all, one cannot but think that the well-being of its intelligent creatures will rank high in the value-scheme of a benign creator. As should really be the case in general, approaches based on the study of nature and the reflections of theology can here be brought into alignment.[12]

12 As regards the Catholic ramifications of the issue, it is certainly true that the Church emphasizes the distinction between body and soul, and views the former, soul, not as a product of the physical causality of nature, but of a special act of creation on the part of God. But this of course need not (and indeed should not) be construed as requiring an unbridgeable gap between doctrine and evolution, since there simply is no need to claim that biological evolution in and of its self creates souls rather than saying that it affords fitting occasions for a choice creation of souls.

However, optimalism as such is compatible with either of these approaches. Both axiological emergence and divine selection – can be contemplated. And either can in theory provide a pathway to the other. In principle, a naturalistic optimalism can regard the institution of an optimalistic mode of things as only natural and to be expected relative to the existence of a benign creator.

A distinction here is that between issues settled by a body of fact and those that are left open by it. If I tell you that there is someone in the room next door, the question "Is the room next door occupied?" is settled. The information at hand *demands* an affirmative answer. But the question "Is there a woman in the room next door?" is left open. The information at hand *permits* an affirmative answer but does not require it. And just this is the situation between being intelligently designed and being the product of an intelligent design. For the fact is that intelligent design as such is neutral, indecisive, and agnostic with respect to theology. It certainly does not authorize the step from "*could* be the product of an intelligent agency" to "*could only* be the product of intelligent agency".

As long as there is a prospect that intelligence can figure as a pivotal factor in cosmic and on biological evaluation, there need be no inherent conflict between noophelic optimalism and enlightened theism: the two points of view need not be seen as reciprocally antagonistic.

12 Historical contextualization

The ideas that the arrangements of the cosmos pivot on interests of intelligent beings in general goes back at least to the Church Fathers.[13] The interests of intelligence – the well-being and thriving of intelligent beings is plausibly seen as the crux for the assessment of ontological merit. To say that one order of things is superior to another is to say that it better serves the interests of intelligence and intelligent beings in the cosmos. The value at issue here with "being for the best" is a matter of being so as intelligent creatures see it – that is from the vantage point of intelligence itself.

And this present line of deliberation has deep antecedence in the history of philosophy. For its answer to the Ultimate Question of existence rests on two lines of deliberation: (1) Axiogenesis: the Optimality Principle of actualization, and (2) Noophelia: the Intelligence-Oriented Standard of valuation. The former is part of the heritage of Leibnizian optimalism. And the second is part of the heritage of

13 See, in particular, the discussion of Origin in Billicsich (1952-59), pp. 217–222.

Plato and Neo-Platonism. The ontology of these present deliberations might be called Axiological Optimalism, with its optimalism deriving from Leibniz and its axiology deriving from Plato.[14]

13 Logical determinism and Burley's principle

Once the existence of a contingent realm is accepted, we immediately arrive at the questions: Why *this* realm? If its reality is indeed contingent, then there will be alternatives. So why is it that existence is as is: why not have some alternative stand in its place? Why are things as is rather than otherwise? And here we come to the question of determinism.

Determinism is the doctrine that everything that is and that happens in the world has to be exactly as is: that reality just does not admit of any unrealized alternatives. There are three main versions of this doctrine:

- *logical determinism*: that the uniqueness of existence roots in that reality as it is does not admit of any coherently characterizable alternatives.
- *physical* or *causal* (*sometimes also scientific*) *determinism*. That everything about the world is the result of the working of inexorable causes.
- *theological determinism*: that reality is and must be as is owing to the unalterable degrees of God (or Fate).

The present deliberations will focus on the first of these. For running through the history or philosophy like a recurrent Leitmotiv is the idea that fundamental logico-ontological considerations that necessitate the existing order of things.

Already Aristotle grappled with the idea and the discussion of logical determinism originates in Chapter 9 of his *On Interpretation*. In his sea-battle discussion there Aristotle conditions (and ultimately rejects) an argument which goes essentially as follows:

14 Leibnizian neo-Platonism is still very much alive. Its axiological approach to ultimate explanation has been advocated in recent decades especialy by two philosophers working along kindred albeit somewhat different lines – John Leslie and the present author: Leslie's relevant publications included: Leslie (1970, 1978, 1979, 1980, 1982, 1989). And the present author's relevant publications include: Rescher (1984, 2000a,b, 2010). Two recent anthologies of relevantly instructive material are: Wippel (1911) and Leslie and Kuhne (2013). Both of these books provide ample further bibliographic information.

(1) All well-defined propositions have a fixed truth status: either true or false.

(2) The truth-value of a well-defined proposition is fixed: it does not change over time.

(3) Since (by (2)) well-defined propositions about future occurrences have a truth-status independently of time, those that concern future events are true (or false) atemporally – and thereby also in absence of the fact. So everything that will happen is already determined.

Aristotle proposes to avert this determinism by rejecting (1).[15] But for those (numerous) logician who regard the (1)-correlative "Law of Bivalence" as correct, the pathway to a version of logical determinism remains wide-open.

Among the medieval schoolmen too some held that the world's facts are as they have to be – that the truth about reality are necessitated by logico-conceptual considerations.

Thus in his influential *Treatise on Obligations*[16] the medieval scholastic philosopher Walter Burley (ca. 1275 to ca. 1345) laid down the rule – let us call it Burley's Principle – to the effect that: *When a false contingent proposition is posited, one can prove any false proposition that is compatible with it*. His reasoning ran as follows. Let the facts be that:

(P) You are not in Rome.

(Q) You are not a bishop.

And now, of course, also:

(R) You are not in Rome or you are a bishop. (P or not-Q)

All of these, so we suppose, are true. Let us now posit by way of a (false) supposition that:

Not-(P)You are in Rome.

15 For further literature see the chapter in "Truth and Necessity in Temporal Perspectives" in Rescher (1969). On the historical aspect of the issue see also Chapter 15: "Determinism and Indeterminism" in Hankisson (1999).

16 Translated in part in Kretzmann and Stump (1988), see pp. 389–412.

Obviously (P) must now be abandoned – "by hypothesis". But nevertheless from (R) and not-(P) we obtain:

You are a bishop. (Not-Q)

And in view of thesis (Q) this is, of course, false. We have thus obtained not-Q cogent inference from by acknowledged truths – *where* Q *is an arbitrary true proposition*. And it is clear that this situation obtains in general. For let p and q be any two (arbitrary but nonequivalent) facts. Then all of the following facts will also of course obtain: $\neg(\neg p)$, $p \wedge q$, $p \vee q$, $p \vee \neg q \vee r$, $\neg p \vee q$, $\neg(\neg p \vee q)$, etc. Let us focus upon just three of these available facts:

(1) p

(2) q

(3) $\neg(\neg p \wedge q)$ or equivalently $p \vee \neg q$

Now let it be that you are going to suppose not-p. Then of course you must remove (1) from the list of accepted facts and substitute:

(1′) $\neg p$

But there is now no stopping. For together with (3) this new item at once yields $\neg q$, contrary to (2). Thus this supposition of ours which runs contrary to accepted fact (viz., not-p) has the awkward and unacceptable consequence that *any other arbitrary truth must also be abandoned*. The supposition that things might be different self-distincts in a *reductio ad absurdum*.

It should be stressed that Burley's Principle is a matter of the *logical* rather than merely *causal* coherence of fact. No doubt the causal fabric of the world is woven pretty tight, and causal connectors ramify across vast reaches of space and time. The world causal fabric is in many ways chaotic with minute changes in one locale makes for big differences elsewhere. (Meteorologists envision a "butterfly effect" in weather forecasting, where the flutter of a butterfly's wings in Sydney can possibly make for a dramatic difference in New York's weather a week later.) But Burley's Principle aims at something quite different; it roots to the logical rather than causal interconnectedness of the manifold of truth.

14 Spinoza and Leibniz

A portentous crossroad in the history of logical determinism came in the dialectical clash between Leibniz and Spinoza.

For Spinoza, God or Nature is the all-encompassing Reality which by its very nature necessitates the totality of all things, seeing that "From the necessity of the infinite things in infinite ways – everything that falls within the scope of the infinite intellect" (*Ethics*, Pt. 1, prop, 16). And in consequence "Nothing in Nature is contingent, but all things are determined the necessity of the divine nature to exist and to act in a particular way" (*Ethics*, Pt. I, prop. 29). Accordingly, "things could not have been produced by God in any other way or in any other order different from that actually produced (*Ethics*, Pt. I, prop. 33). Reality is a manifold in which nothing could possibly differ from what it actually is. Every fact about the world (nature, reality) necessary: nothing is contingent: "A thing is deemed 'contingent' for no reason apart from the deficiency of our knowledge" (*Ethics*, Pt. I, prop 33, scholium 1).

Nothing about the world would possibly be different in any regard for what is it, "For if it were, then God's decrees would have to be different from what he has in fact decreed regarding Nature and its order" and this would mean attributing to God a different intellect and a different will" which would be absurd (*Ethics*, Pt. I, prop 33, scholium 2). And so "All things follow from God's eternal decree with the same necessity by which it follows from the essence of a triangle that its three angels equal two right angles" (*Ethics*, Pt. I, prop 49, scholium, ad fin).

The metaphysics of this mode of logical determinism was revolutionized by a line of thought introduced by Leibniz. His transformative idea was that the concept of necessitation is equivocal in admitting two crucially distinct versions

- *unconditional*, *absolute*, *categorical* (or "metaphysical") *necessity*. The necessity that something must be without any presuppositions or preconditions.
- *conditional* or *unrealized necessity*. The necessity that something must be if some fundamental precondition is obtained.

The distinction at issue tracks the medieval distinction between consequent and consequential necessity, between *necessitas consequentiae* ($p \rightarrow \Box q$) and *necessitas consequentis* $\Box(p \rightarrow q)$. With Leibniz the distinction is implemented via the difference between that which is "metaphysically necessary" (such as the truths of logic and pure mathematics), and that which is "morally" necessary "given the creative actions of a benign God. And this distinction enabled Leibniz to avoid Spinozatic necessitarian by holding that the real world is only one among many possibilities: its existence is not absolutely (metaphysically) necessary but neces-

sary only "morally" (i.e., evaluatively) relative to the creative choice of a benevolent deity.

What Leibniz does, however, concede to the tradition of logical determinism is the reciprocal co-necessitation of the world's facts along the lines of Burley's Principle. For with regard to *conditional* necessitation Leibniz is prepared to hold that given that anything in the world is as is, everything else must be so. The make-up of the world's constituents will not be absolutely (categorically) necessary, but it will be coordinatively co-necessitated. As members of *this* world nothing admits to any change whatsoever: everything must be as is.

Accordingly, Leibniz insisted that there is a crucial difference between saying, with Spinoza, that

– there is nothing *about* the world that could possibly be different from what it is

and saying, as he himself was prepared to do, that

– There is nothing *within* the world that could possibly be different from what it is.

For the former is a contention of absolute necessity, the latter a contention of conditional necessity: *given* that we are dealing with this world, this is how things have to be: change any one thing and you step away from this world and are no longer talking about its members. The former excludes the prospect of other world alternative to this world: the latter excluded the prospect of following different things into this world.

The former excludes the prospect of replacing this world by another; the latter excludes the prospect of making changes within *this* world. The one blocks mind-looking world *replacements*, the other blocks mind-oriented world *changes*. The two positions are radically different.

Leibniz's conceptual innovation made possible the accommodation of seemingly irreconcilable positions. In deflecting all the arguments for logical determinism to the perspective of world internality it left a wide open door to contingency that its acceptance of the externalized perspective of manifold of possible alternatives TO (albeit not IN) the existing world.

As Leibniz saw it, the things of this world – indeed of any possible world – are defined and determined by the properties they have (by their "complete individual concept"). One cannot alter these properties – change them in any respect and you are no longer dealing with the same thing but with something else. (Recall here the discussion of Alexander in Section 8 of the *Discourse on Metaphysics* and that of Caesar in Section 12.)

For Leibniz, there is a crucial distinction between the absolute necessity of "X must be" and the relative or contextual necessitation of "X must be if Y does". The things of this world are contextually neecssitated both as to their being and as to their descriptive nature: each must exist and be as is in the context of any and every other thing. But this mode of co-necessitation is crucially different from the absolute and unconditional necessitation of Spinoza.

With regard to contingency, then, Leibniz occupies a halfway position. An inner or local determinism prevails with respect to any possible world: within any such world everything is necessarily and unattainably as it is. At this local level, a single logical determinism prevails, with no change possible. But at the external or global level of possible worlds at large a pervasive contingency indeed obtains. The existence of any particular world-as-a-whole is entirely optional for the deity (creator). While there is no prospect of change within the world – the entire totality of what is real could indeed be something else. At this global level, contingency is a real and inevitable prospect for Leibniz. But there is a crucial distinction between contingency *in* the world and contingency *of* the world. For him the real world is contingent – it is but one alternative among others. But there is no contingency in the world. Once you enter into it everything is determined to be as it is via the laws of nature decreed by God. Leibniz held that the world's contents are unmodifiable, their make-up cannot be tinkered with. But they could in theory be replaced by others: contingency prevails at the global level. Yet nevertheless the world's things are *co-necessitated*. Each carries all the others in its conceptual nature – remove one from the world and the others cannot remain behind unchanged. Each thing becomes the descriptive hallmark of the entire world to which it belongs and carries all the other substances of its world along as part of its own nature. With any change everything changes and the whole world and all its contents are afforded thereby.

Just here lay the pivot of Leibniz's rejection of Spinoza. For despite the fact that Leibniz too was a necessitarian of sorts, he emphatically rejected.

Spinoza's monolithic position that equated the "necessity" of the world's arrangements with the necessity of mathematical fact. For Spinoza has it that "All things follow from God's eternal decree with the same necessity by which it follows from the essence of a triangle that its three angles equal two right angles" (*Ethics*, Pt. II, prop. 49, ad fin). And Leibniz emphatically rejected this thesis of necessity uniformism.

Leibniz agreed with Spinoza that the nature of any substance – actual or merely possible – is necessitated – that its make-up is unalterable and that if it were to change in any respect it would no longer be the thing it is. All this is part

Display 1
WORLD CHANGE ONTOLOGY

	World-internal Replacement	*World-Alternativeness*
I	Possible	Possible (Radical Indeterminism)
II	Possible	Impossible (Leibniz)
III	Impossible	Possible (Limited Indeterminism)
IV	Impossible	Impossible (Spinoza)

of his doctrine of the "complete individual concept" of things and its associated thesis of the identity of evidential and correlative diversity of the distinguishable. But alike things are inevitably unchangeable they are certainly in principle replaceable – be they microcosmic or cosmic in scope for there will, in theory exist a different universe altogether. (Real existence is never a descriptive feature of a thing encompassed in its "complete individual concept" – a point on which Leibniz and Kant are in complete agreement.) For Spinoza by contrast there is only one manifold of possibility and its existence and nature are alike of invariable necessity.

And so, as Leibniz saw it, Spinoza's position was predicated on a fatal flaw – a failure to recognize a pivotal conceptual distinction – amely that between world-internally local and rationale necessity on the one hand and global or external necessity on the other. As far as Leibniz was concerned, Spinoza was into metaphysical error by the fundamentally logical insufficiency – the failure to recognize the distinction between two entirely different modes of necessitation – the one internal, local and conditional, and the other external, global, and absolute.

This conceptual distinction leads to a quadruplicate taxonomy of possible positions as per Display 1. As this perspective makes clear, Leibniz occupies a halfway house between the radical determinism of Spinoza and the radical indeterminism of an Epicurean position that pure chance prevents all.

15 Kant's Copernican revolution

Spinoza's doctrine of rational determinism was committed to the idea of the inevitability of the real, the idea the abstract general principles of reason ensure that Reality has to be as is and cannot possibly be any different. It was just this thesis that Immanuel Kant regarded as constituting the crux and center of the old "dogmatic" philosophy that he was determined to overthrow in his "Copernican Revolution". After all the *critical* aspect of Kant's *Critique of Pure Reason* was aimed at establishing the limited capacity of human reason to answer our fundamental questions. For, so Kant insisted, the nature of reality is an issue which rationality

as such predicates: all that it can establish is how *we humans must think* of Reality. The only necessity accessible to our intellect is that of how we do and must think of things the questions of necessity as such is beyond our range and for aught that we can ever know about it, reality may well be a matter of contingencies through and through. Kant firmly positioned ultimate questions about the nature of things beyond the reach of the human intellect, in a region any step we take into a quicksand of contradiction.

In taking this sceptical turning Kant was no child of an Age of Reason, Rejecting cognitive optimism regarding the powers of reason, he gave way to a deep pessimism about its incapacity:

> Human reason has this peculiar fate that in one species of its knowledge it is burdened by questions which, as prescribed by the very nature of reason itself, it is not able to ignore, but which, as transcending all its powers, it is unable to answer.[17]

As regards Kant's position on the classic metaphysical question of an ultimate explanation for the nature of things, this doleful declaration says it all.

However, the present deliberations suggest that the source of reason's dissatisfaction lies in asking the right question wrongly by insisting on getting an answer of the wrong kind: one given in the terms of reference of factual rather than evaluative deliberation.

16 McTaggart

The principle defender of logical determinism in modern time was J. M. E. McTaggart. It is of interest to see how his positon fits into the framework of Display 1.[18]

One of the most notable theses of McTaggart's system is his contention that – on grounds of fundamental logical principle – gives any of the world existents everything else in the world must of necessity be just exactly as it is, that given that any substance in the world is what it in fact is, every substance must be just as it is. And so "all that exists, both substances and [their] characteristics, are held together in one system of extrinsic determination" [NE, § 138]. For "every fact about every other substance [within the universe] extrinsically determines every fact about the universe, and ... every fact about the universe extrinsically determines every fact about every other substance" [NE, § 137]. And, therefore, "the

17 Kant (1998), Preface (to the first edition).

18 On McTaggart's metaphysics see Sharma (2015).

supposition that anything should be different from what it is, therefore, is one we have no right to make" [NE, § 139]. The world's things are bound up in one vast network of mutual necessitation.

McTaggart substantiates this position as follows:

> If one of the qualities of [a substance] *A* were not there, there would be no ground to assert that the other would be there. ... If any part of the nature of *A* goes, the nature of *A* as a whole goes. The substance which replaces *A* might have some qualities in common with *A*, just as any other substance might have qualities in common with *A*. But we have no right to subtract *Z* and then say "because we have only subtracted *Z* and because there is no intrinsic determination of *Z* by *X* and *Y*, therefore *X* and *Y* remain". By subtracting *Z*, we have destroyed *A*, and *X* and *Y* were here only as parts of the nature of *A*. No quality of a substance, therefore, could be different while leaving the others unchanged, and no quality of a substance is completely contingent to any of its other qualities. (NE, § 109)

To be sure, McTaggart does not hold categorically that the totality of existing things must be as is, but only conditionally that if something belongs to this totality then it must be exactly as is. And this is so because

> Tt's qualities are part of the nature of the substance, not something apart and distinct from that nature. It is therefore impossible to distinguish the substance from its qualities in such a way as to allow the substances to be different while their natures are the same. (NE, § 149)

Thus McTaggart does not maintain the absolute necessity of the world's existence, but only the relative co-necessitation of its existents. His position is in this regard effectively ideological with that not of Spinoza but of Leibniz.

And so in the end, McTaggart's position is akin to Leibniz's in that his acceptance of reality determinism based on co-necessitation still leaves room for the question of why Reality-as-a-whole is as it is.

But just here Leibniz – unlike McTaggart – has the resources for a resolution readily at hand. For at this point the principle of Optimality can be pressed into service once more. The question "Why is the real world as is rather that otherwise" can again be developed on the basis of a reply that takes the line that "this is so because it too is for the best", And if at this point there comes the question: "But what explains the Principle of Optimality itself? Why should it be that this principle is available to do this requisite work?" Here too a reply is available to the effect that this is the case because it too is for the best. After all, any genuinely ultimate principle of explanation will have to be self-explanatory and cannot avoid being so.

Of course this still leaves open the questions not of the now-resolves question of exploration but rather of the yet open question of motivation: Why is it that we should *accept* the Principle of Optimality and make use of it in addressing our ultimate questions.

This is clearly a large and problematic issue. And it was in order to address *this* issue that Leibniz wrote his *Theodicy*,[19] with its key thesis that if the nature of things has a rational explanation at all – and there must surely be one – then it has to proceed in the order of value (which can be self-substantiating) rather than the order of facts (which cannot be so).

17 Optimalism again

Leibnizian optimalism has many theoretical advantages. Here is just one of them. It is possible – and perhaps even plausible – to contend that *world-existence* – that is to say, the existence of a world – is necessary while nevertheless this world's *nature* (and therewith the existence of this particular world) is contingent. But this would mean that separate and potentially different answers would have to be provided for the two questions "Why is there anything at all?" and "Why is the character of existence as is – why is it that this particular world exists?" For these of course are very different questions. Consider, for analogy, the distinction between "Why is there currently a president of the USA?" (A: The constitution requires and provides for it) and "Why is there currently this particular president of the USA?" (A: the people elected him in 2012.) Such different questions call for different answers.

And so the existence of *some* contingent order of things could well be necessary, even though there is no particular contingent order of things – not even this actual one – whose existence is necessary.[20]

However, an axiogenetic approach enjoys the advantage of rational economy in that it proceeds uniformly here. It provides a single uniform rationale for both answers – namely: "Because this is for the best." It accordingly also enjoys the significant merit of providing self-supportively for the rational economy of explanatory principles at the level of metaphysical fundamentals.

19 On relevant aspects of Leibniz's position see the chapter on "Leibniz and the World's Improvability" in Rescher (2013).

20 There is reason to think that this was in fact the line taken by Leibniz in maintaining the necessity of God's nature as a creator, notwithstanding the optimality-dependent contingency of this particular world of his creation.

Bibliography

Billicsich, F. (1952-59), *Das Problem des Übels in der Philosophie des Abendlandes*, 3 vols., Wein: A. Sexl.

Doyle, A. C. (1892), *The Adventures of the Beryl Coronet*.

Gale, R. M. (1991), *On the Nature and Existence of God*, Cambridge: Cambridge University Press.

Hankisson, R. J. (1999), "Determinism and Indeterminism", in *Cambridge History of Hellenistic Philosophy*, edited by K. Algra, J. Barnes, J. Mansfeld, and M. Schofield, Cambridge: University of Cambridge Press.

Kant, I. (ed.) (1998), *Critique of Pure Reason*, edited by P. Guyer, and A. W. Wood, Cambridge University Press.

Kretzmann, N. and Stump, E. (eds.) (1988), *he Cambridge Translation of Medieval Philosophical Texts, Vol. I: Logic and Philosophy of Language*, Cambridge: Cambridge University Press.

Leibniz, G. W. (1969), *Philosophical Papers and Letters*, translated and edited by L. E. Loemker, Dordrecht: Reidel.

Leslie, J. (1970), "The Theory that the World Exists Because It Should", in *American Philosophical Quarterly*: 7, 286–298.

Leslie, J. (1978), "Efforts to Explain All Existence", in *Mind*: 87, 181–197.

Leslie, J. (1979), *Value and Existence*, Oxford: Clarendon Press.

Leslie, J. (1980), "The World's Necessary Existence", in *International Journal for Philosophy of Religion*: 18, 207–223.

Leslie, J. (1982), "Anthropic Principle, World Ensemble, Design", in *American Philosophical Quarterly*: 19, 141–151.

Leslie, J. (1989), *Universes*, London and New York: Routledge.

Leslie, J. and Kuhne, R. L. (eds.) (2013), *The Mystery of Existence: Why Is There Anything At All?* , Oxford: Wiley-Blackwell.

Pruss, A. R. (1988), "The Hume-Edwards Principle and the Cosmological Argument", in *International Journal for Philosophy of Religion*: 434, 149–165.

Rescher, N. (1969), *Essays in Philosophical Analysis*, Pittsburgh: University of Pittsburgh Press.

Rescher, N. (1984), *The Riddle of Existence*, Lanham MD: University Press of America.

Rescher, N. (2000a), "Price of an Ultimate Theory", in *Philosophia Naturalis*: 37, 1–20.

Rescher, N. (2000b), *Nature and Understanding*, Oxford: Clarendon Press.

Rescher, N. (2010), *Axiogenesis: An Essay on Metaphysical Optimalism* , Lanham MD: Lexington Books.

Rescher, N. (2011), "On the Improvability of the World", in *The Review of Metaphysics*: 64, 489–514.

Rescher, N. (2013), *On Leibniz*, Pittsburgh: University of Pittsburgh Press.

Rowe, W. L. (1970), "Two Criticism of the Cosmological Argument", in *The Monist*: 54, 441–459. Reprinted in *Philosophy of Religion: Selective Readings*, 2nd edition, edited by W. L. Rowe and W. Wainwright, New York: Harcourt Brace Jovanavich, 1989, 142–156.

Rowe, W. L. (1975), *The Cosmological Argument*, Princeton: Princeton University Press.

Sharma, R. K. (2015), *J. M. E. McTaggart: Substance, Self, and Immortality*, Lanham, MDL: Lexington Books.

Van Inwagen, P. (2015), "Nothing is Impossible", in *God, Truth, and Other Enigmas*, edited by Miroslaw Szatkowski, Berlin: De Gruyter, 33–58.
Wippel, J. F. (ed.) (1911), *The Ultimate Why Questions: Why Is There Anything at All?*, Washington DC: Catholic University of American Press.

Chris Daly

Van Inwagen on Testimony and Contingency

Abstract: This paper assesses two arguments that Peter van Inwagen has offered concerning rational belief revision. The arguments purport to show that we should radically decrease our degree of belief in philosophical propositions given certain relatively uncontroversial data. One argument concerns philosophical testimony against a philosophical proposition that you believe. The other concerns the contingency with which you have certain philosophical beliefs rather than others.

1 Introduction

This paper assesses two arguments that Peter van Inwagen has offered concerning rational belief revision. The arguments purport to show that we should radically decrease our degree of belief in philosophical propositions given certain relatively uncontroversial data. Call the arguments 'the argument from expert testimony' and 'the argument from contingency'.

2 The argument from expert testimony

This argument concerns the topic of disagreement between equally informed and rational agents (epistemic peers). In recent years the literature on this topic has enormously blossomed, but I want to confine my attention to van Inwagen's treatment of it.[1] This is partly in order to keep my discussion manageable but, more importantly, it recognizes van Inwagen's pioneering role in raising the topic and putting a new philosophical problem squarely on the philosophical agenda.

Epistemic peers with respect to some topic are people who match with respect to their training, reasoning power, memory, level of attention, and the evidence available to them. It is an idealization to take there to be any epistemic peers.[2]

1 I will set aside, for instance, Thomas Kelly's influential work. Against van Inwagen, Kelly has recently claimed that 'there is *no* plausible view about the epistemology of disagreement, on which philosophical agnosticism is compelling' (Kelly (2016), p. 375).

2 King (2011) argues that peer disagreement is rare and that we rarely have reason to count people who dissent from us as epistemic peers. My concern, however, is with disagreement with epi-

https://doi.org/10.1515/9783110664812-022

More realistically, what we have, with regard to any topic, are near epistemic peers – people who do not markedly differ in the above cognitive and evidential respects. For any philosophical topic, there are near epistemic peers who disagree. Even more strikingly, for many, if not all, philosophical topics on which you have a view, there are epistemic superiors who disagree with you. An epistemic superior in philosophy is someone who has more philosophical ability than you. David Lewis has more philosophical ability than you if he has more understanding – for every distinction, theory and argument you understand, he understands them and more – and he has more creativity – he formulates more (and more illuminating) distinctions and theories than you.

The opinions of epistemic superiors raise a puzzle: what should you think if an epistemic superior about a certain topic disagrees with you about that topic? Van Inwagen has made this issue vivid in the following way (Van Inwagen (1996), p. 138, and Van Inwagen (2010), pp. 23–24). Suppose you reject compatibilism in the free will debate because you are convinced by the consequence argument: the argument that says that if your actions are due to laws of nature acting on events which happened centuries ago, and since you are not responsible for those laws or events obtaining, then as a consequence your actions are not free. Now, David Lewis has more philosophical ability than you and he understands the consequence argument at least as well as you. There is no reasoning in the argument that you follow which Lewis fails to follow and there are no distinctions drawn in the argument that Lewis fails to observe. Yet for all that he rejects the argument. How should you respond? Should you remain as confident as before about the argument, or should your degree of confidence in it decrease somewhat, or should you follow Lewis and reject the argument altogether? It would seem partial to retain your original degree of confidence in the argument. The fact that Lewis rejects it is evidence that the argument is unsound.

By the same reckoning, it would seem partial to suppose that you have a privileged philosophical insight into the argument. (van Inwagen has tentatively suggested that he has insight into the truth of incompatibilism that Lewis lacked: see his Van Inwagen (1996), pp.30, 34, and 40–43, and Van Inwagen (2010), p. 26. By 'insight' van Inwagen means evidence that he possesses that he cannot communicate). The fact that Lewis rejects the argument is evidence that you lack such insight (or evidence that if you have 'insight', your insight is mistaken). Of

stemic superiors. King's points do not bear on this issue. Whether or not such disagreement is extensive, it is surely more frequent than peer disagreement: David Lewis held some philosophical views each of which put him in a small minority of believers yet David Lewis was the epistemic superior of most philosophers on most philosophical topics.

course Lewis might be smart without being right, and you might not be as smart as him but still be right. But why suppose that it is you rather than him who possesses insight other than the fact that you believe the consequence argument and he doesn't? If insight tracks smartness, the odds are that Lewis has insight. If insight does not track smartness, who knows what it might track? It's then merely even odds that you have it. (Van Inwagen (2010), p. 27, records that he finds his own suggestion hard to believe).

What should we conclude? There seem to be three different conclusions that could be drawn that challenge our prior convictions. All of these conclusions state that it is mistaken to hold fast – i.e., to fail to adjust credences to bring them closer to the credences of another party. The weakest conclusion is that, while still believing the consequence argument is sound, your degree of belief should be decreased somewhat. A stronger conclusion is that you should suspend judgement and become agnostic about the consequence argument. The strongest conclusion is that you come to share Lewis' degree of belief in the consequence argument – i.e., you reject the consequence argument. I take it that Van Inwagen resists all of these conclusions: he records that he 'is unable to give up his beliefs in many of these cases [such as the consequence argument] and unable to accept the conclusion that his own beliefs are not rational, but is also unable to answer satisfactorily the arguments for the skeptical view' (Van Inwagen (2010), p. 10). Certainly van Inwagen resists the second and third conclusion. When he says that he is 'unable to give up his beliefs in many of these cases', I take it that he is unable even to decrease his degree of belief in each of these cases. For my part, I endorse the weakest conclusion.

I will address two responses to van Inwagen's puzzle before offering my own. The first response would be to look in askance at this entire issue. According to this response, the emphasis on testimony and adjusting your credence in the light of expert opinion seems off the mark. It is doubtful, the response continues, whether these factors play an important role in the more theoretical side of science either. And in both philosophy and in science, the premium is on understanding why, not taking something for granted because someone, however clever, tells you so, or even being more inclined to believe it on those grounds.

In reply, I agree that in philosophy the premium is on understanding why. But let's see how this plays out. Lewis offers an argument for p. I think I understand the argument but I don't accept p. What to do? Should I take it that I fully understand the argument and am justified in not accepting p? Or should I take it that I don't fully understand his argument and that I'm not justified in not accepting p? Here's a consideration in favour of the second option: Lewis is a smarter philosopher than I am. He understands any argument better than I do. If there was anything wrong with the argument, he'd have spotted it. But he offers the

argument. So it's a good argument. I should then take myself not to have fully understood the argument. I should take it that he does, however, and, since he accepts its conclusion, so should I.

I grant that testimony doesn't play an important role in the theoretical side of science, and perhaps it plays only a small role in philosophy as well. But these points do not carry us very far in addressing van Inwagen's puzzle. What is it for something to have an important role in a discipline? Distinguish the issue of how extensively testimony is used in making claims in science – which we will suppose is small – from the issue of what epistemic warrant testimony provides claims in science – which seems to be an independent issue. Consider a parallel: mathematics makes only infrequent use of experiments – such as using computers in non-deductive reasoning to reach certain mathematical results (Baker (2008). Note that Baker argues that the use of computers and non-deductive reasoning are symptomatic rather than defining features of experimental mathematics). But this leaves open the question of what degree of warrant such experiments confer on those claims: do they strongly confirm those claims, weakly confirm them, or lend no warrant at all? In similar fashion acknowledging that testimony is widely employed in neither philosophy nor science (or so we will suppose) does not address the question of how much warrant a given piece of testimony would confer on a given claim. But this is only to say that the response under consideration does not address van Inwagen's puzzle.

Incidentally, the comparison of philosophy with science is pertinent for another reason. Van Inwagen's puzzle arises in *any* discipline, including any scientific discipline (as van Inwagen recognises). Consider the epistemic situation of Georges Lemaître in the late 1920s. Lemaître calculated that the universe is expanding – a then novel and striking result. The two most eminent physicists of the day, Eddington and Einstein, were both apprised of Lemaître's work but dismissive of his bold claim (Farrell (2005)). Given the latters' expertise and Lemaître's juniority – he didn't even have his PhD when he proposed his hypothesis – how could he be rational in retaining his belief? The fact that Eddington and Einstein disagreed with Lemaître surely had some epistemic weight, but how much? In particular, did it have enough weight to make it irrational for Lemaître to retain his belief in his claim?

A second response is that van Inwagen's conclusion is defeatist – it counsels us, the small fry of philosophy, to abandon the subject to the big players – and how can that be right? In reply, we should distinguish epistemic from motivational issues. The issue highlighted by van Inwagen's puzzle is epistemic: whether you are entitled to remain as confident as you are in your philosophical opinions given that you have an epistemic superior who disagrees with you. This differs from the motivational issue of whether you should persist in doing philosophy given

that you have an epistemic superior who disagrees with you. Perhaps in that case you are like a sparring partner for a world boxing champion: although you know you're going to lose every contest, you also know that you're performing a useful service by keeping the champion and yourself in good shape. Or perhaps you have other forms of motivation – incessant curiosity, a need for intellectual stimulation, and the like. No matter. It can be reasonable that your motivation for doing philosophy remains undiminished despite the fact that you know you have epistemic superiors who disagree with you. (Notice too that, so far as the motivational issue goes, your epistemic superior need not disagree with you about this issue). The epistemic issue, however, persists. Even if you have undiminished reason to do philosophy despite the fact that an epistemic superior disagrees with you, what should the degree of confidence of your philosophical opinions be given that you have an epistemic superior who disagrees with those opinions?

Very well, let's review our options. There seem to be three of them: (1) Abandon belief that p. (2) Continue to believe that p but concede that one is not justified in believing that p. (3) Continue to believe that p, believe that one is overall justified in believing that p, but concede that one's belief that p rests on assumptions that others reject and for which one has no knockdown argument.

Van Inwagen rejects option (2) and I agree. Consider, then, option (1). Should you follow Lewis and reject the consequence argument? Here are some considerations against this. Lewis thinks the consequence argument is flawed and he even wrote a paper to show this (Lewis (1981)). If Lewis were both your epistemic superior and right, he would be able to show epistemically inferior philosophers such as you and I what is wrong with the consequence argument. He would be able to construct a sound counter-argument, explaining each step of it in such careful and accessible detail that even much more limited philosophers such as ourselves would be able to follow it and become convinced. But this is not what we find: when we read what he has to say, we understand it but we come away unconvinced. So Lewis is not both your epistemic superior and right. Now, he is assuredly your epistemic superior. Therefore, there is some reason to think that he is not right in rejecting the consequence argument. So you should not take option (1).

I will now present my case for option (3). A distinction can be drawn between two kinds of testimony. In a case of *partial disclosure* I know that a certain person is my epistemic superior and that that person believes that p. But that is all the information I have. In particular, in this case I do not know what that person's reasons for believing that p are. By contrast, in a case of *full disclosure*, I know

that a certain person is my epistemic superior, that that person believes that p, and what their reasons are for believing that p.

I suggest that cases of partial disclosure may carry more epistemic weight – provide more reason to believe that p – than cases of full disclosure *in so far as expert testimony provides reason to believe a claim*. Suppose I accept incompatibilism but I read, say in a blog, that David Lewis has a forthcoming paper in which he criticizes incompatibilism. The blog does not say what Lewis's criticisms are. Given that I recognize David Lewis as my epistemic superior on this issue (as in so many others), I thereby have good reason to decrease my degree of belief in incompatibilism. Call the fact that Lewis is my epistemic superior and that he rejects incompatibilism, *the testimonial fact*. A case of the foregoing sort is an illustration of partial disclosure. It is a case in which all that is disclosed is the testimonial fact. Now, suppose instead that not only do I learn that Lewis has a forthcoming paper that criticizes incompatibilism, I also get to see a preprint of the paper and I understand the criticisms he offers. This is an illustration of full disclosure: in addition to Lewis's testimony, I get to understand his reasons for his so testifying. Notice, however, that in this case when I think about whether or not I should continue to believe in incompatibilism, the facts that Lewis is my epistemic superior and that Lewis has testified against incompatibilism seem subsidiary to the facts about how good Lewis's reasons are rejecting incompatibilism. If my assessment is that Lewis's reasons are good ones, I thereby have good reason to reject incompatibilism. The testimonial fact will still give me reason to reject incompatibilism but then it will be far from my strongest reason to reject incompatibilism. On the other hand, if my assessment is that Lewis's reasons are poor ones, this undermines whatever reason is conferred by the testimonial fact. Suppose Lewis's reasons are poor ones, or at least I would judge them to be poor reasons if I were apprised of them. In that case, partial disclosure – my simply knowing the testimonial fact – would give me more reason to reject incompatibilism than full disclosure would.

The case van Inwagen describes is one of full disclosure: we are fully informed about Lewis's reasons for rejecting the consequence argument. Now, since Lewis sets out his reasons for rejecting incompatibilism, we should look to them and assess them. If we are unpersuaded by Lewis's reasons – as presumably van Inwagen is – we should reject his case for incompatibilism. But, by the same token, we then have defeaters for the testimonial fact. So our degree of belief in incompatibilism should not be lessened.

Here is a final consideration. Even if Lewis were the most epistemically superior philosopher around, many other leading philosophers disagree with his assessment of the consequence argument (such as van Inwagen himself!). I suggest that these philosophers are collectively an epistemic match with Lewis at least

with respect to this issue.[3] So what should you think about the consequence argument in the light of this? You are caught between conflicting sources of testimony, conflicting epistemic superiors. And it is likely that you will find yourself in this predicament for every philosophical view you hold. What should you do? You could become an agnostic about every philosophical issue (or deny that these are genuine issues of disagreement) and become an agnostic about almost every philosophical argument. But that does not help. You would still be disagreeing with your epistemic superiors since they think that you should not be agnostic but should be agreeing with them.

Also that suggestion seems to assign too much weight to the views of others in philosophy, even if they are your epistemic superiors. Bill Lycan endorses the following Principle of Humility: 'If you maintain that *p*, but others who are as intelligent and perceptive as you and have given as much consideration to the matter and are not biased in any nonepistemic way deny that *p* on the basis of reasons, you do not know that *p*' (Lycan (2013), p. 117). Lycan's principle concerns knowledge – something we haven't been addressing here. That aside, his principle needs qualifying. Suppose I have epistemic peers who deny *p* on the basis of reasons. But suppose also I have good reason to think that defeating or undermining conditions for those people's reasons obtain. Then it seems I can still be in a position to know *p*. In the case of the above option of becoming agnostic, that option seems to overlook the evidence and the arguments that you possess and the justification that they confer on your view, despite the fact that you know that some other philosophers disagree. They may disagree with you and they may be your epistemic superiors but that does not license your abandoning your position anymore than the ancient Greeks would have been licensed in disbelieving what their eyes told them about motion because Zeno was their epistemic superior in philosophy. Philosophical testimony has some weight, but it can be outweighed by what you perceive or by what you argue. (Chalmers (2015), 14 takes a similar line).

Van Inwagen would presumably take me to be thereby committed to endorsing the following speech, where *p* represents the name of some 'substantive philosophical doctrine':

> I accept *p*, and I regard all trained philosophers who are in my epistemic position (that is, who are aware of the arguments and other philosophical considerations I am aware of) and who accept the denial of *p* as irrational. Such philosophers are in violation of their epistemic obligations. They are comparable to ordinary, educated people of the present day who be-

3 *X* is an epistemic match with *Y* with respect to some topic *S* iff *X* is not *Y*'s epistemic superior with respect to *S* and *X* is not *Y*'s epistemic inferior with respect to *S*.

> lieve that cigarette smoking does not cause lung cancer or that the positions of the stars and planets at the moment of one's birth determine one's fate. (Van Inwagen (2010), p. 19)

I endorse the first part of this speech. If I think that the balance of argument favours *p*, then I should think that those who have the same data (the same evidence, the same principles; in short, the same total evidence) as I do but who fail to believe that *p* have committed some error in reasoning. That is, I should think that they are irrational: they fail to believe what the balance of argument enjoins. If the data is consistent and entails *p*, but some philosophers fail to believe *p* while accepting all of the data, then they are in error. Or if they accept all of the data, the best explanation of the data is *p*, they have considered *p*, but they fail to believe *p*, then they are in error. In either case, those philosophers are irrational in accepting the data but in rejecting *p*. I do not, however, endorse the second part of the speech. The error (or errors) made by philosophers who think that the balance of argument does not favour *p* are no doubt subtle ones, since weighing up the data and the arguments involved in a philosophical debate is a difficult and involved process. (Van Inwagen would agree: Van Inwagen (2010), p. 21. As he has said in another context, philosophy is argument without end).

Van Inwagen notes that some philosophers may suffer from cognitive disadvantages of one sort or another (such as having less philosophical talent than oneself) and suggest that that 'excuses them of the charge of irrationality' due to the 'ought implies can' principle (Van Inwagen (2010), p. 23). To my mind, however, that does not so much excuse them of irrationality as *explain* their irrationality – it explains why they reject a proposition which follows from, or which best explains, the data available to them. On a narrowly interpretation of 'can', these philosophers individually cannot draw the conclusion that *p* from the data: they do not have the intellectual wherewithal to do so. But on a broader interpretation of 'can', they can draw this conclusion. A reading group of philosophers each of whom has mediocre philosophical talent might pool their intellectual resources and collectively draw a valuable and novel conclusion that none of them individually would be able to reach. Having reached this stage, each of the philosophers concerned would be in a position to review the group's reasoning to draw the conclusion themselves. A single mediocre philosopher might be guided in drawing a conclusion, one that he could not otherwise draw, with deft coaxing and assistance from a more talented philosopher – the Socratic method in operation. These cases illustrate how, in a broader sense of 'can', these philosophers can conclude that *p* and so the 'ought implies can' principle is not contravened.

This distinction between a narrow and a broad sense of 'can' is not contrived. It is a distinction that we need to mark in any case. Consider a moral example: suppose I am a reckless gambler and have blown all my family and friends' savings

at the tables. I ought to repay them. In a narrow sense of 'can', however, I cannot do so because I haven't the money. In a broader sense, though, I can do so. I can work all the hours available to me and for however long it takes for me to repay my debts. And it is this broader sense of 'can' that is operative in applying the 'ought implies can' principle to this example.

In any case, contrary to the passage quoted from van Inwagen, there is no comparison between philosophical disagreement with astrology or smoking. For each of the latter the negative case has been clearly and decisively made (decisive by the standards of science and ordinary thinking) and it has been extensively disseminated. For these reasons, and unlike error in these other matters, philosophical error is understandable and excusable.

My take on option (3) does not permit one to shirk the epistemic obligation of weighing up all the evidence and all the arguments concerning whether p. What it does permit is that, having weighed up all the evidence and argument, one believes that p even if one cannot provide others with an argument that is compelling, by their lights, that p. Is this being closed minded and dogmatic? I would have thought that to be closed minded or dogmatic would be to believe that p without taking stock of all the evidence and all the arguments, or to be selective or arbitrary in ones' assessment of these. That is not the case here. Instead, what we have here is a belief that p that is, by one's own lights, overall justified on the basis of all the evidence and argument, despite the fact that other philosophers, including those more expert than one, have examined the same evidence and argument but fail to believe that p.[4]

3 The argument from contingency

Van Inwagen presents a different argument in his paper 'Freedom to Break the Laws' but again the target is the credence we assign to philosophical claims. Van Inwagen supposes that he has perhaps reached equilibrium between his philosophical theories and his data, but it strikes him that his data could so easily have been different:

> The point of the philosophical equilibrium I occupy depends (perhaps) on predispositions to belief inherent in my genes, (very likely) on what my parents taught me about morals and

4 Bricker (2006), pp. 68–69, defends a similar view which he calls 'epistemic chauvinism': that your belief that p may be knowledge that p even if someone with the same evidence, the same conceptual competence and the same capacity for reasoning believes that not-p.

> politics and religion when I was a child, and (certainly) on what university I selected for graduate study in philosophy, who my departmental colleagues have been, the books and essays I have read and haven't read, the conversations I have had at APA divisional meetings as a result of turning right rather than left when I was wandering aimlessly about at a reception. … Other philosophers have reached different points of philosophical equilibrium simply because these factors have operated differently in the course of the formation of their opinions. These reflections suggest – and the suggestion is very strong indeed – that I ought to withdraw from the point of philosophical equilibrium I occupy and become a sceptic about the answers to all or almost all philosophical questions. (Van Inwagen (2004), p. 342)

Notice that the above argument applies as much to each of us as it applies to van Inwagen, and that it applies to our epistemic peers and superiors alike. Or perhaps we should say 'so-called epistemic superiors'. For if the argument goes through, no philosopher is epistemically superior to you (and none epistemically inferior to you). For all of these philosophers formed their opinions in highly contingent circumstances of the kind van Inwagen describes. But what the argument concludes about the opinions of van Inwagen can also be concluded, mutatis mutandis, about the opinions of all other philosophers: those opinions ought to be withdrawn as unwarranted.

It is a striking feature of the argument's conclusion that no matter how much evidence you have as compared with another philosopher, and no matter how many sources it is drawn from, and how much longer and harder you've thought about the topic, your philosophical opinions are no more warranted than theirs. This is virulent scepticism.

I agree with van Inwagen's reflections but disagree with his conclusion. First, the phrase 'what my parents taught me about morals and politics and religion' in the quoted passage should alert us that the reflections he entertains do not apply only to philosophical views: they apply to pretty much of all our views. (Perhaps some of our beliefs are innate and so are not products of our current environment and education.) They are reflections on the contingent circumstances in which we form most of our beliefs. If you had been born in North Korea or in an Amish community or in Saudi Arabia, you would have grown up with very different views about politics, culture, science, religion and history than the ones that you do have. It's a sobering thought about what might have been and about the dependence of which beliefs we have on how our lives have happened to turn out, but it doesn't single out our philosophical beliefs for special consideration.

Second, let's turn specifically to the case of philosophy. On the basis of the evidence available to me, and assuming that I'm at least moderately rational, I have formed certain philosophical views. I realize that many other philosophers have different views and I realize that they have formed them on the basis of differ-

ent evidence.[5] I do not know what their different evidence is: I have not been privy to their experiences, chance meetings and reflections. And if I did have the evidence they have, perhaps I would not have formed the beliefs that I did or perhaps I would not have formed those beliefs with the degree of belief that I did. What am I do to? The rational thing to do is to form my beliefs on the evidence I have and to be prepared to update my beliefs when I fall into conversation with these other philosophers and learn about their reasons for the different beliefs that they hold. But, *ex hypothesi*, I have already formed my philosophical beliefs on the basis of the evidence, however partial, available to me (where this evidence includes the evidence that others have evidence for different beliefs to mine.) And, similarly, since I am moderately rational, I will seek out further evidence and be prepared to update my beliefs when new evidence comes in. With all that in place, there is nothing else for the rational person to do and there seems to be no special problem concerning my philosophical beliefs.

It is very salutary to note that I could have had, and that others do have, different bodies of evidence or different methods of evidence-gathering or of reasoning. This may prompt me to query whether my actual evidence is extensive, whether it is representative of the total available evidence (much of which I might not have), and whether my methods of evidence gathering and of reasoning are reliable (King (2011), pp 267f). I may have methods and techniques for testing these assumptions of mine, but then presumably van Inwagen would think that much the same issue arises in turn for them. Whatever methods or techniques I appeal to would be open to the same challenge, and so it is a challenge that cannot be met in these terms.

What is at issue here is the nature and distribution of epistemic luck. There is an extensive literature on this topic. (Pritchard (2007) provides an overview). For our purposes in discussing van Inwagen's view, the key point is the following. Certain kinds of epistemic luck undermine our putative knowledge and thereby undermine our beliefs but (it is widely agreed) other kinds do not. For example, the luck of existing or of being suitably equipped to form the rele vant beliefs does not undermine your knowledge that you exist or that there is a barn in front of you. The puzzle about van Inwagen's argument is that it is undiscriminating between these two kinds of case. Never mind where we think philosophical beliefs fall – whether they fall on the undermined side of the distinction or not. The considerations van Inwagen offers apply to all of our beliefs indiscriminately: as we have seen, if his argument against philosophical beliefs is successful, it ramifies and all

5 Or perhaps they have the same evidence as I do but have made a different assessment of the evidence. I will take this option as read in what follows.

beliefs alike are undermined. That proves too much. And for all that van Inwagen has said, it is an open question which side of the undermined/non-undermined distinction philosophical beliefs would fall on even if we were to introduce into the discussion facts that discriminate between beliefs falling on different sides of the distinction. In particular, so far as van Inwagen's argument goes, it is an open question whether, in such circumstances, they would without exception fall on the undermined side of the distinction. For instance, it is plausible that if we cannot provide a reasonable epistemology for claims about some field, *X*, then any beliefs that we have about that field are undermined. Accordingly, if a reasonable epistemology of our philosophical beliefs was impossible – if there could be no reasonable account of how our philosophical beliefs are reliable – then those beliefs would be undermined. But such a contention goes beyond the datum given to us in the discussion, the datum that it is highly contingent which philosophical beliefs each of us has.[6] Finally, there does seem something to the idea that it is not reasonable to maintain a given belief once you think that there is no reason to suppose that whatever evidence you have supports that belief (Feldman (2006)). But, as I see it, that is not the case van Inwagen presented. He outlined counterfactual situations in which you would have had different philosophical beliefs than the ones you actually do because of the different training, information and the like that you would have had in those circumstances. The fact that, in different circumstances you would have had different evidence, and so formed different beliefs does not seem to indicate that you lack reason to believe that the evidence you actually have supports the beliefs you actually have.

4 Conclusion

Let me summarize my overall line of argument by setting out some Socratic questions – a complex of alternative lines of Socratic questioning laid out in a sort of flowchart. (I take this device from the closing pages of Van Inwagen (1996)).

Do you think that the fact that your philosophical beliefs are formed in highly contingent circumstances and that you might easily not have had whatever reasons you have for holding them undermines your beliefs? If so, does this apply to your belief that: the fact that all of your beliefs are formed in highly contingent

6 Incidentally, I wonder how the challenge raised by the contention could be met. Any reasonable epistemology of *X* will require a philosophical theory about *X* – a theory about the nature of truths concerning *X*. But belief in any such theory will automatically be subject to the same challenge. How could the challenge be met without begging the question against it?

circumstances and that you might easily not have had whatever reasons you have for holding them undermines your beliefs? Indeed, does it apply to all of your beliefs? Moreover, do you think that anyone else is better epistemically placed than you? If so, why is this? What bestows epistemic privilege on them? And why can't you acquire that privilege if you don't already have it? If not, then shouldn't you think that you have no epistemic superiors (or epistemic inferiors)?

Well, do you think that some philosophers are epistemically superior to you and that their testimony is a reason for changing your beliefs? If not, game over. If so, and you recognize that David Lewis is an epistemic superior of yours who rejects your view that p, then do you revise your belief that p simply because of the foregoing fact about Lewis? If so, what do you say about the fact that other epistemic superiors of yours disagree with Lewis and accept p? If not, then do you revise your belief that p because of the foregoing fact about Lewis *and* the fact that you understand the reasons Lewis has given for rejecting p? If not, why do you revise your belief that p? If so, do you find Lewis's reasons persuasive? If so, why does the fact that *Lewis* provides those reasons enter into your thinking about whether or not p? Why should it matter who provides those reasons? Shouldn't you just be evaluating those reasons on their own merits? If not – if you don't find Lewis's reasons persuasive – then why does the fact that *Lewis* provides those reasons enter into your thinking about whether or not p? To repeat: why should it matter who provides the reasons? And if you think that Lewis's reasons are not good ones, then why don't you think that he is less your epistemic superior than you previously thought that he was? In fact, since you disagree with him, and if you think that his reasons for rejecting p are not good ones, why don't you think that you've got the better of him with respect to this dispute?

Bibliography

Baker, A. (2008), "Experimental Mathematics", in *Erkenntnis*: 68, 331–344.

Bricker, P. (2006), "Absolute Actuality and the Plurality of Worlds", in *Philosophical Perspectives*: 20, 41–76.

Chalmers, D. J. (2015), "Why Isn't there More Progress in Philosophy", in *Philosophy*: 90, 3–31.

Farrell, J. (2005), *The Day Without Yesterday: Lemaître, Einstein, and the Birth of Modern Cosmology*, New York City: Basic Books.

Feldman, R. (2006), "Epistemological Puzzles about Disagreement", in *Epistemology Futures*, edited by S. Hetherington, Oxford: Oxford University Press, 216–236.

Kelly, T. (2016), "Disagreement in Philosophy: Its Epistemic Significance", in *The Oxford Handbook of Philosophical Methodology*, edited by H. Cappelen, T. Szabó Gendler and J. Hawthorne, Oxford: Oxford University Press, 374–394.

King, N. L. (2011), "Disagreement: What's The Problem? Or A Good Peer is Hard to Find", in *Philosophy and Phenomenological Research*: 85, 249–272.
Lewis, D. (1981), "Are We Free to Break the Laws?", in *Theoria*: 47, 113–121.
Lycan, W. G. (2013), "On Two Main Themes in Gutting's What Philosophers Know", in *The Southern Journal of Philosophy*: 51, 112–120.
Pritchard, D. (2007), *Epistemic Luck*, Oxford: Oxford University Press.
Van Inwagen, T. (1996), "Is it Wrong Everywhere, Always, and for Anyone to Believe Anything on Insufficient Evidence?", in *Faith, Freedom and Rationality*, edited by J. Jordan and D. Howard-Snyder, Savage, Maryland: Rowman and Littlefield, 137–154.
Van Inwagen, T. (2004), "Freedom to Break the Laws", in *Midwest Studies in Philosophy*: 28, 334–350.
Van Inwagen, T. (2010), "We're Right. They're Wrong", in *Disagreement*, edited by R. Feldman and T. A. Warfield, Oxford: Oxford University Press, 10–28.

Eleonore Stump

The Problem of Evil and Atonement

Abstract: It is widely supposed that the doctrine of the atonement is the distinctive doctrine of Christianity. In this paper, I focus on the confluence of the doctrine of the atonement and theodicy. There is reason to expect that there would be some connection between the two On traditional Christian attempts at theodicy, something about suffering somehow conduces to the ultimate good for human beings, which is union with God. But the atonement is also supposed to provide this ultimate good. And so it seems that there should be some intrinsic connection between the benefits for human beings postulated by the doctrine of the atonement and the benefits supposed to justify God's allowing suffering in the world. In this paper, I explore the nature of this connection.

1 Introduction

In the course of his long and distinguished career, Peter van Inwagen has written extensively about the problem of evil, carving out a position that has been widely influential and is still the subject of much discussion today.[1] Although my own views on the problem of evil diverge from van Inwagen's at various points, I have learned a great deal from his work on this subject.[2] And, in fact, underlying the disagreements, there is considerable agreement about some fundamental claims. Among them are the claims that the ultimate good for human beings is union with God, that this good is imperiled by post-Fall human tendencies towards evil, and that suffering, generically considered, has some role to play in promoting that ultimate good.[3]

In this paper, I want to move forward from these general points of agreement by exploring in a preliminary way the connection between attempted solutions to the problem of evil and the doctrine of the atonement,[4] specifically the view that

1 See Van Inwagen (1988a,b, 1991, 1996, 1997, 2000, 2004, 2005, 2006).

2 The theodicy that I myself favor I have explained and defended, at length, in Stump (2010).

3 In the convergence on these fundamental points, of course, there is nothing distinctive about my views and van Inwagen's. These claims have been espoused by the overwhelming majority of Christian thinkers from the beginning of the Christian era to now.

4 For a more detailed explanation of the basic idea of atonement, see Stump (2012a).

https://doi.org/10.1515/9783110664812-023

there are redemptive effects for human beings stemming from the passion and death of Christ.[5]

It is widely supposed, by both Christians and non-Christians alike, that the doctrine of the atonement is the distinctive doctrine of Christianity; and Christians often speak of the value of the atonement itself as infinite, or so great as to be incommensurate with all other created goods.[6] Of course, the doctrine of the atonement has been interpreted in varying ways in the history of Christian thought.[7] And, traditionally, the passion and death of Christ have been thought to have not just one effect but several. (Aquinas, for example, says in addition to its redemptive effects, Christ's passion operated as a source of merit, as a sacrifice, and as satisfaction for human sins.[8]) Here, I will focus on the doctrine of the atonement with regard to the redemptive effects of Christ's passion and death; and I will try to do so in a way that sidesteps differences among the varying interpretations of the doctrine and that leaves aside considerations of any other effects of Christ's passion and death.

My purpose in restricting the focus in this way is to explore the confluence of the redemptive effects of the atonement and the role of suffering in God's providential plan for bringing human beings to union with God.[9] There is reason to expect that there would be some confluence between the doctrine of the atonement and a solution to the problem of evil (assuming that there is one). In particular, we could reasonably expect that in a consistent Christian theology there would be some connection between the redemptive effects of the atonement, on the one hand, and the good promoted by human suffering, on the other.

On one of the traditional Christian claims van Inwagen and I both accept, something about suffering somehow conduces to the ultimate good for human beings, which is union with God.[10] But the atonement is also supposed to provide

5 Or, the incarnation, passion, death, and resurrection of Christ. Here I will focus on the passion and death of Christ, but the main points would not be altered if the focus included the incarnation and resurrection as well.

6 See, for example, Plantinga (2004). For some critical commentary on the theodicy Plantinga proposes in this paper, see McCord (2008). I share some of her concerns about attempts at theodicy based largely or wholly on comparisons of the summed value of worlds.

7 I have discussed some of these in Stump (2012a) and Stump (2013b).

8 Cf., e.g., ST III. 48.

9 And even here I am limiting myself to those salvific effects that have to do with what I have elsewhere called 'the future problem of sin', namely, the sinful dispositions in a human psyche. (See Chapter 15 of Stump (2003)). There is also a past problem of sin, which I am leaving out here in the interests of brevity. I am addressing that problem in this context in Stump (forthcoming).

10 See Chapter 13 of Stump (2010) for a defense of the claim that allowing suffering is justified only if the benefit brought about by the suffering goes primarily to the sufferer.

this ultimate good. And so it seems that there should be some intrinsic connection between the benefits for human beings postulated by the doctrine of the atonement and the benefits supposed to justify God's allowing suffering in the world. The doctrine of the atonement and a theologically acceptable explanation of suffering in the world ought to have a connected place in one grand unified theological theory of everything in God's providential plan of salvation.

To begin to explore what this connected place could be, it helps to think about atonement more generally. 'Atonement' is an invented word composed of 'at' and 'one' jammed together with 'ment'. It was once a neologism, coined to express the nature of the solution to a problem. Traditionally, that problem has been understood as the absence of oneness or closeness between God and human beings.[11] However exactly we are to understand closeness and union between persons, it will at least include harmony between their minds and wills.[12] A person who does something that is contrary to the will of a perfectly good God introduces distance between himself and God. Understood in this way, distance between God and a human person has at least a partial source in a human person's failure to will what God wills.[13]

In fact, as Augustine and the Christian tradition after him understood the absence of unity between God and human beings, it is a function of the post-Fall human proneness to sin, where sin is understood as something that is contrary to the will of a perfectly good God.[14] As Augustine himself helped to highlight,[15] one

11 Of course, although this word is relatively new, religious rituals of atonement are old. They figure prominently in the Hebrew Bible, as, for example, in the prescriptions for the day of atonement laid out in Leviticus.

12 For a detailed discussion of the issue, see Chapter 6 of Stump (2010).

13 For discussion of the issue and arguments for this claim, see Stump (2012b).

14 Whether one takes morality to be grounded in God's will or in God's nature, it remains the case that willing what is morally bad is willing something contrary to the will of God.

15 So, for example, commenting on his own failures to will what he himself takes to be the good and wants to will as the good, Augustine says,

> the mind commands the body, and is presently obeyed: the mind commands itself, and is resisted. … it commands that itself would will a thing; … [and it] never would give the command, unless it willed it; yet it does not [will] … [what it] has commanded …. it commands, … [because] it wills: and … the thing [is not] done which … [it] commanded, … [because] it wills it not…. But it does not command fully, therefore is not the thing done, which it commanded. For were the willing full, it never would command it to be, because it would already be.

hallmark of this post-Fall defect in the will is the will's intractability to itself, its proneness to moral wrong even against its own desires for the good.

On this view, human beings will what they take to be good, but they can also simultaneously will against it. Furthermore, although human beings can have a second-order will to will the good, that higher-order will is frequently rendered ineffective by contrary first-order willings of what is not good. A person in such a condition wills the morally wrong thing he himself desires not to will. Worse yet, a human being can be in so lamentable a moral condition that even his second-order will is corrupted. He not only fails to will what is good, but he also fails to will to will what is good.

In addition, it is part of traditional Christian views that a human being cannot be united around what is morally evil.[16] The objective moral standard, at least in its rudiments, is so accessible to ordinary reason that no human intellect is ever totally in ignorance of it. Someone who takes something objectively evil, such as beating his wife, to be morally acceptable or even morally good will also have some awareness, however tacit or covert, that it is in fact morally reprehensible. Any wrong-doing therefore generates a kind of double-mindedness in the wrong-doer. And double-mindedness generates a correspondingly conflicted set of desires and volitions.[17] Consequently, it is not possible for a person's mind or will to be internally integrated in evil; no one is completely single-minded or wholehearted in evil.[18] For these reasons, internal integration is possible only for a person single-mindedly understanding and wholeheartedly desiring the good.

The problem to which the atonement is the solution therefore can be thought of as the human tendency to fragmentation in the will, brought about by the human proneness to sin. And, on orthodox Christian views, the atonement is a sufficient solution for this problem. That is, the atonement offers redemption to all those who will avail themselves of it; it makes union with God available to every-

Augustine, Confessions VIII. 9. I like and have therefore used (with slight modifications) the translation by William Watts, (Cambridge, Mass.: Harvard University Press, 1968).

16 For a discussion and defense of this position, see Stump (2010): Chapters 6 and 7.

17 For a discussion of the connection between intellect and will presupposed here, see Stump (2003).

18 See, for example, ST II-II q.45 a.4 and a.6. The relevant biblical claim can be found in Isaiah 48:22.

one who does not reject it.[19] Somehow,[20] the atonement brings it about that, in the end, those human beings who do not reject the salvation offered in the atonement are united to God. And, in union with God, they are internally integrated around the good. They will only what God wills.

Given this way of understanding the doctrine of the atonement, however, what role is left for suffering in the process of providing the ultimate good for human beings? If the atonement makes union with God available to all human beings, then what happens to the view that suffering also has a role to play in promoting union with God? Is there something central to achieving union with God that is not provided by the atonement but that suffering facilitates? Does suffering somehow enable whatever the means are by which the atonement promotes union with God?

Or put the question the other way around. On the face of it, it seems in principle possible to have a world in which there is the post-Fall human proneness to sin and also the atonement to provide salvation from that post-Fall condition, but in which no one suffers, because the proneness to sin is never successful in bringing harm to others and there is no other source of suffering either. In such a world, what if anything that would conduce to the ultimate good for human beings would be lost by the absence of suffering?

Presenting and defending a grand unified theological theory that brings together the doctrine of the atonement and any kind of attempted solution to the problem of evil is obviously beyond the scope of a short paper. But even a preliminary answer to these questions would give some insight into the connection between the doctrine of the atonement and explanations for God's allowing suffering; and so it would be propaedeutic to such a grand unified theory. In this paper, I want to sketch such a propaedeutic account.

To this end, I will first give a short characterization of the nature of suffering. Then I will highlight the different routes by which suffering can move a person to alter what he wills. Among these routes are two very different ways in which a sufferer can try to harmonize his will with God's will. I will argue that only one of these two ways has any chance of success at bringing the sufferer either to peace

19 These claims are true even for those who accept a doctrine of double predestination. On that doctrine, those who reject the salvation provided by the atonement were predestined by God to reject it. In that sense, the atonement was not intended for them, even if the value of the atonement was in itself so great that it would have sufficed for salvation for everyone.

20 The various interpretations of the doctrine of the atonement differ in their understanding of this 'somehow'. There is no creedal formula that effectively distinguishes among them and privileges just one as orthodox. For some discussion of the main kinds of interpretation in the history of Christian thought, see Stump (2012a).

or to harmony with God's will. Then I will change course and consider justification and sanctification, two processes central to salvation, on orthodox Christian views; and I will examine the way in which Christ's passion and death function to promote each of these processes. Finally, I will return to the paper's main question and outline a way in which both justification and sanctification are promoted by the only promising way of willing what God wills in consequence of suffering.

The result will be an account of the way in which Christ's suffering in his passion and death can be interwoven with the suffering of a person Paula to work her salvation. This account falls short of a full explanation of the confluence between the doctrine of the atonement and an acceptable theodicy (assuming there is one). It is not sufficient for a grand unified theological theory of God's providential plan of salvation. On the contrary, it is only programmatic. But, I hope, it will nonetheless be enough to suggest directions for incorporating the views about the role of suffering in salvation that van Inwagen and I both share into an account of the redemptive effects of the atonement.

2 The nature of suffering

If the problem to which the atonement is the solution consists in the post-Fall human proneness to sin, then overcoming the human failure to will the good is central to the redemptive work of the atonement. Overcoming it requires a person to be integrated around the good. God wills only the good, however; and so a person's proneness to sin and fragmentation in will is healed to the extent to which she wills what God wills. Consequently, if suffering also has a role to play in bringing about the ultimate good for human beings of union with God, as attempted solutions to the problem of evil have typically supposed, then something about suffering too should contribute to bringing a person to the state of willing what God wills.

Now perhaps the most notable characteristic of most suffering is its aversive character. Generally, a person in the midst of suffering is motivated to try to end his suffering;[21] and so suffering has a motive power. One way to explore the role of suffering in the process of bringing a human person to willing what God wills

21 The claim has to be qualified in this way, of course, because there are some kinds of pain or suffering that human beings actually seek out. So-called natural childbirth is one example, and masochistic practices are another. But it is enough for my purposes that most pain and suffering is aversive for most sufferers. For a more detailed discussion of suffering, see Chapter 1 of Stump (2010).

is therefore to consider the differing ways in which the aversiveness of suffering can move a person.[22]

In this regard, it is helpful to recognize that suffering itself is a function of the will. Although it is tempting to think of suffering just as pain, a little reflection shows that this way of characterizing suffering is mistaken. A human being can suffer even without experiencing pain, as, for example, when a person suffers from an unjust political system of which he approves and to which he is strongly committed on ideological grounds.[23] As I have argued elsewhere,[24] a better way of thinking about suffering is to see it as a function of what we care about.

Considered in this way, there are two sides to suffering, an objective side and a subjective side.

Because every human person has some care about what kind of person he is and about his flourishing as that kind of person, part of what it is for him to suffer is for him to be kept, to one degree or another, from flourishing. This is an objective side of suffering, since there is an objective fact of the matter about what will make a human person flourish.

On the other hand, however, what we care about has a subjective side too. This is something to which a person is committed but which is not identical with his flourishing and which may not even be compatible with it. What is at issue in this subjective side of suffering can be thought of as the desires of the heart.[25] When the Psalmist says, "Delight yourself in the Lord, and he will give you the desires of your heart", we all have some idea of what the Psalmist is promising. Suffering also arises when a human being fails to get a desire of his heart or has and then loses a desire of his heart.

Human beings care about two kinds of things, then, their own objective flourishing and also those things that are the subjective desires of the heart. Suffering arises when something impedes or removes either of these kinds of things

22 I say 'can have' here because, of course, the aversiveness of suffering can also drive the sufferer into psychic states that increase the fragmentation in his will and so alienate him from God. Insofar as a human person's will is free, suffering cannot determine any particular psychic state in that person. For a discussion of these issues, see Stump (2010): Chapters 8 and 13.

23 In such a case the diminishment of flourishing can constitute an aversive stimulus even if pain does not.

24 For detailed discussion of this claim, see Stump (2010): Chapter 1.

25 The expression "the desire of the heart" is ambiguous. It can mean either a particular kind of desire or else the thing which is desired in that way. When we say, "the desire of his heart was to be a great musician", the expression refers to a desire; when we say, "In losing her, he lost the desire of his heart", the expression refers to the thing desired. I will not try to sort out this ambiguity; I will simply trust to the context to disambiguate the expression.

that people care about.[26] Another way to put this conclusion is that people suffer because they lose or do not get what they care about.[27] Consequently, as many thinkers in different cultures and times have pointed out, human suffering arises because of human desire.

For this reason, also as noted in the history of philosophy and religion, there are two different ways in which human suffering can be diminished or warded off. Either (1) a human person can bring it about that she gets what she cares about and wills to have, or else (2) she can will to accept what she gets and cease to care about what she does not get.

Sometimes the first of these ways is efficacious. A person exerts himself or is lucky; and the result is that, even in the face of initial setbacks causing suffering, he succeeds after all in getting what he wants. But, of course, cases of this sort are in the minority by comparison with cases in which the sufferer is helpless to change his circumstances and by that means get what he wants.

When a sufferer cannot get what he wants because the circumstances will not bend to his efforts, then if he persists in wanting exactly what he cannot get, he will not only perpetuate his suffering, he will add to it the frustration of his impotent efforts and the anger or anxiety of his checked will. Insofar as suffering is aversive, this sort of reaction to the failure to get what one wants is unstable. A frustrated, angry or anxious sufferer will be motivated by his suffering and by his unhappy reactions to it to seek some more peaceful psychic state.[28]

For these reasons, one might suppose that a sufferer would or should prefer the second way to try to avoid suffering. That is, unless his efforts to get what he wants are successful, the aversiveness of his suffering would or should move him to attempt to will to accept what he gets and to cease to care about what he does not get.

In fact, one might suppose that this is the very reaction to suffering that orthodox Christianity mandates. The ultimate good aimed at by the atonement is union with God, which requires harmonization of the divine and the human wills. But, one might reasonably suppose, on Christian doctrine no instance of suffering

26 To say that suffering is a function of what a person cares about is not to make all suffering a subjective matter. To see the point here, consider Aristotle's claim that all human beings desire happiness. Even though this desire is something subjective, human happiness and the things conducive to it have an objective character, as Aristotle shows.

27 There are, of course, cases in which a person is confused about what would make him flourish, or cases in which what he cares about in regard to flourishing are incompossible things. For complications involving such cases, see Stump (2010): Chapters 1, 13, and 14.

28 Of course, he might not act on that motivation. For some people, there is a kind of consolation in chewing on anger, for example, which they may be loath to give up.

would occur if God willed that it not occur. So it might seem that, from a Christian point of view, any suffering that actually occurs does so only because it is in accord with the will of God. For a person Jerome to will what God wills, then, it might seem that, with regard to either his flourishing or his heart's desires, Jerome must will to accept whatever he gets when something contravenes his flourishing or his heart's desires. Consequently, insofar as willing what God wills is necessary for the ultimate good of union with God, then it appears that the second approach to suffering has to be the right one, on Christian views. A person should let go of his care for his own flourishing or his heart's desires, and he should will to accept what he gets.

3 Willing what God wills

3.1 The no-self view

But this conclusion is hasty. It rests on a particular interpretation of what it is for a human person to will what God wills,[29] and not all the major thinkers in the Christian tradition have accepted it.

One medieval thinker who has accepted it is Meister Eckhart, and it will be helpful to see his view in order to highlight the somewhat complicated alternative to it.

In discussing what he takes to be the right reaction to suffering, Eckhart says,

> Seneca, a pagan philosopher, asks: 'What is the best consolation in sorrow and in misfortune?' And he says: 'It is for a man to accept everything as if he had wished for it and had asked for it'; for you would have wished for it, if you had known that everything happens by God's will, with his will and in his will. … [In] that alone, that it is God's will that it should happen so, a good man's will ought to be so wholly one and united with God's will that he and God have only one will, though that should be for the man's harm or even his damnation. … [A]ll his blessedness consists in … willing and wanting to know nothing but God's will.[30]

There are many other passages in Eckhart's work that one could cite to confirm the impression of his views given by this comment on Seneca's views. Elsewhere, for example, Eckhart says,

29 It also rests on an unsophisticated understanding of what God wills. See the discussion of the difference between the antecedent and the consequent will of God in Chapter 16 of Stump (2003).
30 Eckhart (1978b).

> [T]hrow all anxiety out of your heart, so that in your heart there be nothing but constant joy. … [E]ven if I had to see with my own eyes my father and all my friends killed, my heart would not be moved by it. … I have rightful joy only when neither sufferings nor torments can ravish it from me. Then I am translated into the divine being where no suffering has a place.… If you reach a state where you feel neither suffering nor vexation from whatever may happen, so that suffering is not suffering for you and that all things are sheer joy for you, then the child is truly born [that is, you have achieved spiritual regeneration.] [31]

Or, to take one last example, which makes the point in a more radical way, Eckhart says,

> it is not sufficient for us to have a detached attitude of mind at a specific point in time when we wish to bind God to ourselves, but rather we should have a practiced detachment… We must learn to free ourselves of ourselves…, not holding on to what is our own or seeking anything, either profit, pleasure, inwardness, sweetness, reward, heaven or our own will. God never gives himself, or ever has given himself, to a will that is alien to himself, but only to his own will… [T]he more we cease to be in our own will, the more truly we begin to be in God's will. … We must train ourselves in self-abandonment until we retain nothing of our own.… We should establish ourselves… in the best and most precious will of God through a pure ceasing-to-be of our will and desire.[32]

As Eckhart sees it, then, for a person Jerome to will what God wills, Jerome can have a second-order will to have a will that wills what God wills, but he must not have a first-order will for anything at all. He must not desire, as Eckhart puts it, any profit or pleasure or reward. He must not even desire union with God in heaven. His first-order will must be empty, as it were, in order to enable him to accept whatever happens.

In fact, Eckhart is willing to go even further. In some places, he writes as if he thinks even the desire for willing what God wills has to be stamped out. In explaining the spiritually excellent state of being poor in spirit, he says,

> So long as a man has this particular wish to fulfill the ever beloved will of God… then this man still has a will with which he wants to satisfy God's will… [A]s long as you have the will to fulfill God's will, … you are not poor [in spirit]; for he alone is a poor man who wills nothing and desires nothing. [33]

As Eckhart sees it, then, even if Jerome has no first-order desires for anything in particular, if he still has a second-order willing of his own, then what Jerome is willing is his will, and not God's will. On Eckhart's way of thinking about the mat-

31 Eckhart (1978b).
32 Eckhart (1994).
33 Eckhart (1978a).

ter, the only way *really* to will what God wills is for a person to have no will of his own at all, on either the first-order or the second-order level, so that the will operative in a human person is just the will that is God's will (as distinct from the human will in harmony with God's will).

Since a person's will is at least in part constitutive of his self, however, then if Jerome has no will of his own, even at the second-order level, there is a sense in which Jerome has so self either. In some places, Eckhart writes as if this is just the end he supposes people should seek. So, for example, he says,

> You might ask 'When is the will a right will?' The will is perfect and right when it has no selfhood and when it has gone out of itself, having been taken up and transformed into the will of God.[34]

If these texts represent Eckhart's view accurately, then we would have to say that for Eckhart a person who succeeds in willing what God wills simply has no will of his own at all. The will operative in him is the will that is God's and not his own, human will.

A person in this condition will certainly diminish his suffering. Insofar as suffering results from the loss or undermining of what a person cares about, either as regards his heart's desires or as regards his flourishing, a person who cares about nothing will not suffer either. As Eckhart claims, such a person will have only joy, even if he should see his father being killed before his eyes.

The problem for Eckhart is that the response he advocates to suffering is destructive of the self, and the destruction of the self is itself destructive of the very possibility of union with God.[35] That is because union between persons requires two wills to unite.[36] This fact is the reason why an internally fragmented post-Fall will is an obstacle to the ultimate good for human beings of union with God. There cannot be union between God's will and a human will when the internal divisions in the human will keep that will from being *one* will. By parity of reasoning, there cannot be union between God and a human person if there is no will at all in the human person.

---

34 In Eckhart (1994), O. Davies translated the word 'selfhood' by 'Eigenschaft'.

35 For the distinction between the stamping out of the self and the denial or crucifixion of the self recommended in the New Testament, see Stump (2010): Chapter 14.

36 On orthodox Christian views, there is only one will among the three persons of the Trinity. But the unity among the persons of the Trinity is a metaphysical as-it-were substantial unity and not a coming together into union of two separate entities. For metaphysical unity of the sort at issue in the Trinity, only one will is in fact necessary.

And so in virtue of attempting to stamp out all will, or at least all first-order will, in a human person, Eckhart's recommended response to human suffering makes that response at least as deadly as the original disorder in the will, for which the suffering was supposed somehow to be therapeutic.

Furthermore, there is something inhuman about a person who tries to live out Eckhart's recommended response to suffering. Human beings are social animals, and so part of what it is to be human is to be connected to others in social and caring ways. To care for another person, however, is to desire both the good for that person and union (of one or another appropriate kind) with that person.[37] Eckhart might be able to argue that in desiring nothing for another person, one is somehow willing what God wills for that person and therefore willing the good for that person, although it seems to me that this argument would not be easy to make. But what one cannot imagine is that a person who accepted Eckhart's position could desire any kind of union with another person. Someone who desires another person, as a mother desires her child or a lover desires his beloved, cannot be correctly described as a person who has no desires at all. And so to the extent to which one has stamped out desire in oneself, one has also stamped out the desire for union with others. In this lack of a desire for union with any particular others, there is also a lack of care for them; and that lack of caring connection to others has something inhuman about it.

In addition, because there is something inhuman about it, ordinary human nature will kick back against Eckhart's position. And even for those people who do attempt to adopt it, Eckhart's recommended response is guaranteed to be unsuccessful, for the same reasons that the will cannot be integrated in evil. No matter how much in the grip of Eckhart's theory a person is, some part of his intellect will recognize, however tacitly or covertly, that the destruction of his flourishing or his heart's desires is not good. He will be unable not to see, for example, that the killing of his father before his eyes is not good. And to that same extent he will be unable not to desire that it not happen. So the attempt to stamp out in oneself all first-order desires will result only in the fragmentation of the will that was the problem in the first place. And, mutatis mutandis, similar things can be said about the attempt to stamp out any desire in oneself even at the second-order level, which Eckhart sometimes seems also to recommend.

Finally, insofar as Eckhart's approach to suffering is unsuccessful, it will also not alleviate the aversiveness of a person's suffering.

37 See the discussion of love in Chapter 5 of Stump (2010).

For these reasons, Eckhart's interpretation of willing what God wills and his recommendation of his brand of willing what God wills will be at least as unsuccessful at ameliorating suffering as the attempt to avert suffering by getting what one wills.

3.2 Having a self to deny

Against the background of Eckhart's views, it is easier to understand the different interpretation of willing what God wills that can be found in Aquinas's work. In my view, Aquinas's interpretation and his recommendation of willing what God wills is a much more promising approach to suffering, both as regards Christian theology and as regards the amelioration of suffering. Like Eckhart, Aquinas maintains that the goodness of the human will requires its conformity to the divine will;[38] but his understanding of this conformity is very different from that of Eckhart.

One way to begin to see the difference is to consider Aquinas's response to objections against the claim that human beings are obligated to will what God wills. As Aquinas presents these objections, one objection argues that a human person does not always know what God wills and that this ignorance militates against the obligation to will what God wills.[39] In reply to the objection, Aquinas says,

> we know that whatever God wills, he wills it under the aspect of the good. Consequently, whoever wills a thing under any aspect of the good has a will conformed to the Divine will, as to the reason of the thing willed. But we know not what God wills in particular: and in this respect we are not bound to conform our will to the divine will.[40]

In explaining his reply to the objection, Aquinas says,

> the will tends to its object according as it is proposed by the reason… And therefore if a man's will wills a thing to be according as it appears to be good, his will is good… Now a thing may happen to be good under a particular aspect, and yet not good under a universal aspect…and therefore it comes to pass that a certain will is good from willing something considered under a particular aspect, which thing God wills not [insofar as it is considered] under a universal aspect…[41]

38 ST I-II q.19 a.9.

39 ST I-II q.19 a.10 obj.1.

40 ST I-II q.19 a.10 ad 1. The translation in this and the following quotations is from the standard translation of the Fathers of the English Dominican Province, which I like and therefore have used, though sometimes with slight modifications.

41 ST I-II q.19 a.10.

On this way of thinking about harmony between God's will and a human will, there is harmony when, like God, a human being wills the good because it is good. And this harmony can exist even when the particular things wanted by the divine and the human will are different. God grasps the good universally, that is, with all things considered, whereas a human mind is limited in its apprehension of the good. So God can reject as not the all-things considered good something that a human being rightly takes as good and wills because it is good; and yet it can nonetheless be true that the human will is willing what God wills, in virtue of willing the good because it is good.

Aquinas thinks that this distinction between willing the good under a particular aspect and willing it under a universal aspect contains the answer to yet another objection to the claim that a person is obligated to will what God wills. Aquinas puts that objection this way:

> if a man were to will what God wills, this would sometimes be contrary to filial piety: for instance, when God wills the death of a father: if his son were to will it also, it would be against filial piety.[42]

Here Aquinas is addressing the sort of case discussed by Eckhart. Eckhart's view is that a person Jerome trying to will what God wills should accept the death of his father as a joyful thing because it is in accordance with God's will; and so, from Eckhart's point of view, the objection Aquinas is considering is just mistaken. From Eckhart's point of view, willing that his father die when his father does die is actually required for Jerome to will rightly. But, as Aquinas sees it, the objection is correct in supposing that it would be contrary to right willing if Jerome were to will the death of his father when Jerome's father does in fact die. And yet, Aquinas thinks, the objection is confused in supposing that in order for Jerome to conform his will to God's will, Jerome needs to will the death of his father if this is what God wills.

On Aquinas's view, Jerome's will can be conformed to God's will even when Jerome's will and God's will are in opposition as regards the death of Jerome's father.[43] To be in conformity with God's will, Aquinas holds, it is enough for

42 ST I-II q. 19 a.10 Obj.3.

43 Aquinas's position here depends crucially on his distinction between God's antecedent and God's consequent will. Roughly put, God's antecedent will is what God would have willed if everything in the world had been up to him alone. God's consequent will is what God actually does will, given what God's creatures will. For Aquinas, the will with which God assents to suffering is only his consequent will, not his antecedent will. For a discussion of this distinction with regard to the problem of suffering, see Chapter 13 of Stump (2010).

Jerome to will what he wills under the aspect of the good, for the sake of the good. To will something in this way is to will what God wills, namely, the good, and also to will it as God wills, namely, out of love of the good, that is, out of charity.[44] For Aquinas, then, a person Jerome can will the opposite of what God wills – say, with regard to the death of Jerome's father – and still be willing what God wills, as long as Jerome is willing what he rightly takes to be a real good and willing it because it is good.[45]

For Aquinas, the model for willing what God wills and in fact the model for the right response to suffering,[46] is given by Jesus in his three-fold prayer in the Garden of Gethsemane.[47] In the first part of the prayer, Jesus asks that God would let the cup pass from him. It is evident that in this part of the prayer Jesus is giving voice to his desire not to die. So, (in his human will[48]) Jesus is opposed to his death; his will does not accept it as a good thing. On the other hand, Jesus finishes the prayer by saying to God, "not my will but yours be done".

In commenting on this prayer, Aquinas holds that in his prayer Jesus willed something different from what God willed, because Jesus willed not to die when God willed that Jesus die. Nonetheless, Aquinas thinks, in willing not to die, Jesus' will was still in conformity to God's will.[49]

Using Aquinas's interpretation of willing what God wills, described above, we can gain insight into Aquinas's thought here by understanding Jesus' will during his prayer in terms of the hierarchical structure of the will.[50] On this understanding, (in his human nature) Jesus had a first-order desire not to die. In this desire, he was rightly taking not dying as a good and desiring it because it is good. But he also understood that God's grasp of the good could be different from his own, so

44 ST I-II q.19 a.10.

45 More nuance is needed here in order to make this position consistent with Aquinas's view that an erring conscience binds. In cases where the conscience of a person Jerome is in error, then on Aquinas's views Jerome will be double-minded about the goodness of what he is willing, and he will not be whole-hearted in willing it either. But these nuances, important as they are for other purposes, do not need to be pursued for my purposes here.

46 See, for example, ST III q.21 a.3 and a.4, where Aquinas explains the different ways in which the prayer in the Garden of Gethsemane instructs human beings and gives them a model to follow.

47 See Luke 22:42, Mark 14:36, and Matthew 26:39. For the remaining prayers in Gethsemane, see Matthew 26:42-44.

48 This qualification is needed because, on the Chalcedonian formula for the incarnate Christ, there are two wills in Christ. One of these is the human will of Christ, and the other is the will of God. Insofar as the will of God cannot be opposed to the will of God, the only will of Christ that can be in any sense different from the will of God is the human will of Christ.

49 ST III q.18 a.5.

50 For more discussion of this issue and Aquinas's position, see Chapter 9 in Stump (2003).

that his dying might be the good universally considered, as Aquinas puts it. And so (in his human will) with regard to his dying, Jesus had a second-order desire for a will that wills what God wills.

Even though his first-order desire was in opposition to God's will as regards his dying, Jesus' first-order desire was in harmony with God's will insofar as it willed what it willed as a good and willed it out of love of the good. In addition, Jesus' second-order desire was straightforwardly in harmony with God's will insofar as Jesus wanted to will what God willed, even if God's will should turn out to be in opposition to Jesus' first-order desires.

As Aquinas sees it, then, in this prayer in Gethsemane, (in his human will) Jesus has a first-order desire for what he himself really does want, namely, not to die, and a second-order volition[51] for a will that wills what God wills, even if God's will is that Jesus die. Both the first-order desire and the second-order volition are part of Jesus' will in this prayer.

As is made clear by the last part of prayer he prays in Gethsemane and the rest of the story, after the first part of his prayer, Jesus comes somehow to understand that God's will is that Jesus die then. Because Jesus has a second-order volition for a will that wills what God wills, his second-order volition combines with his understanding of God's will to elicit in him a first-order desire to die then. But this new first-order desire, to die then, does not stamp out the desire not to die. He does not go to his death joyously, for example. And so the two opposed first-order desires exist in him simultaneously.

Consequently, insofar as God's will is that Jesus die then, and Jesus comes to believe that it is, his first-order will is divided against itself. But because Jesus is committed to the desirability of letting God's desires take precedence over his own, the first-order desire in himself which becomes the volition on which Jesus acts is the first-order desire to die then. In this rank-ordering of God's will over his own first-order will not to die then, however, Jesus does not give up his desire not to die. He still has that desire; he just acts counter to it, because his volition to will what God wills is effective in determining which of his divided first-order desires he acts on. In this sense, Jesus exhibits both his persistent maintaining of his own

51 As I am using the term 'volition' here, a volition is a desire that would be effectual in bringing about the thing it wills if nothing external to the will impeded it. When an agent's second-order desire counts as a volition, it is because that second-order desire is effectual in bringing the agent's first-order desires into harmony with it, so that the first-order desire which becomes the agent's first-order volition is the first-order desire that the agent wants to have. For more discussion of this distinction between desire, generically considered, and volition as a species of desire, see my Stump (1988).

self (contrary to the sort of position Eckhart recommends) and the denial of that self in the interest of willing the good that God wills.

And so, as Aquinas sees it, in his prayers in the Garden of Gethsemane Jesus in effect teaches people both that they should will what God wills and that they can count as willing what God wills even when they desire something that God does not will.[52]

On Aquinas's view, therefore, the right response to suffering is complex. It requires a person Paula to continue to hold on to the desire for the thing that is lost or withheld in the suffering; but at the same time it requires her to accept the loss because she desires to have a will that wills what God wills. That is, she desires that her father not die, even while she accepts the death of her father in virtue of willing to will what God wills. This complex state in her will is matched by a complex state in her intellect. She takes the death of her father to be in some sense a bad thing, but she also takes as a good thing accepting this bad thing, in virtue of the fact that she takes God and God's will to be good and she recognizes the death of her father as God's will.

If one thinks in terms of medical care, the difference between Eckhart's view and Aquinas's view may become clearer. Suppose that Paula's father is being subjected to a painful bone marrow transplant in the hospital. Then in the grip of a spirit such as Eckhart's, Paula might simply accept with joy her father's having this procedure, on the grounds that it is what the doctors want. Or, in the Thomistic spirit, Paula might both want and not want her father's having this procedure. She might take the medical things done to her father as bad, even while she accepts that agreeing to them is good because it is in accordance with what the doctors want for her father. In this condition, Paula's desire that her father not have the transplant is overridden but not extirpated by her desire to will what the doctors will. And Paula has this second-order will because she takes the doctors and the will of the doctors to be good.

This complex state in the will in response to suffering can provide some of the amelioration of suffering sought by the acceptance Eckhart advocates, but it does so without the failure and inhumanity attaching to Eckhart's approach. The persistence of the thoroughly human first-order desires even in the face of suffering provides the humanity of the Thomistic approach, and not having to try to stamp out those first-order desires adds to the chance of success for those trying to adopt Aquinas's approach. On the other hand, the second-order desire to will what God wills governs the will's first-order desires and so produces some unity in the will.

52 ST III q.21 a.2.

The sufferer manages at the same time to have a will of her own and to ally her will with God's will in her suffering.

In what follows, for ease of reference, I will call this complex first- and second-order state in the will 'willing what God wills in the Thomistic sense', or just 'the Thomistic will' for short. In the Thomistic will, there is a surrender of the self (because of the second-order desire for willing what God wills even when it is contrary to one's own first-order will) but not an abandonment of the self (because of the persisting first-order desire for what is lost in suffering). Furthermore, although the Thomistic will is still an internally divided state, it does have an integrated character. The will is not at war with itself, because the second-order volition rules the first-order desires. And insofar as the integration is around willing what God wills because God and God's will is good, the integration is an integration around the good.

In suffering, then, a person can simply war against his suffering and rail against the forces that produced it; but by doing so he will exacerbate his suffering with the pains of frustration and anger or anxiety, and he will feel that he does. But, in the nature of things, the aversiveness of suffering will move a person to try to find some diminishing of the suffering. Under the pressure of his suffering, a person can try Eckhart's approach. He can try to give up what he cares about and instead find acceptable whatever happens to him. But, as I have argued, this move is self-destructive and likely to fail. On the other hand, if he foregoes Eckhart's option, suffering can move him in the direction of the Thomistic will. If he forms the Thomistic will, then he will come to some degree of peace, because there is some acceptance of his suffering, but without the sacrifice of his humanity, since he does not give up what he cares about.

To the extent to which the sufferer finds a place of peace, there will be for him some amelioration of his suffering; and this place of peace will come when he is willing what God wills, in the Thomistic sense. In this condition, and to the extent to which the sufferer wills what God wills, the sufferer is also integrated around the good.

So one thing suffering can do is move a person in the direction of peace, which comes with that integration around the good that is possible to have in a post-Fall world, as modeled by Christ's prayer in Gethsemane.

4 Suffering and atonement

No one should mistake these thoughts about suffering for a theodicy, even if they do elucidate one claim central to some theodicies, namely, that suffering has a

role to play in remedying the post-Fall human proclivity to fragmentation in the will brought about by sinfulness. A theodicy attempts to provide a justification for God's allowing suffering; and some good stemming from suffering is not by itself sufficient to justify anyone's inflicting or allowing that suffering. In any event, my concern here is not with theodicy. It is rather with the exploration of a connection between the atonement and the role a theodicy can assign suffering in the process of salvation.

The first thing to notice in this regard is that there are two different ways in which any such connection might be drawn.

On the one hand, we could suppose that something about Christ's passion and death is the primary provider of the means to union with God, and we could ask how the Thomistic will elicited by suffering contributes to the efficacy of Christ's passion and death. In this case, the role of suffering in evoking the Thomistic will would be instrumental to the work of the atonement.

On the other hand, however, it is not out of the question, that is, it is not incompatible with any orthodox Christian doctrine, that the atonement provides its redemptive effects at least in part by synergistically enhancing suffering's ability to elicit the Thomistic will from a sufferer. Seen this way, for at least some part of the process of salvation, the primary work of salvation would be done by human suffering; and Christ's passion and death would somehow enhance or enable that work. In that case, for that part of the process of salvation, the atonement would be instrumental to the eliciting of the Thomistic will through suffering.

In my view, with the elucidation of the Thomistic will given above, it is possible to discern a progression in which *both* of these connections have a place.

4.1 The Thomistic will and operative grace

To see this progression, it is important to focus on one component of the Thomistic will, namely, the second-order desire for a will to will what God wills. Where does this second-order will come from?

On the account I have been arguing for above, when a person comes to the point of framing his will into the configuration of the Thomistic will, the aversiveness of suffering has played a major role in getting him there. But, on orthodox Christian views, this role of suffering cannot be the whole story of the genesis of the Thomistic will. The second-order desire in the Thomistic will is a will to will what God wills; and since what God wills is only the good, this second-order desire can also be understood as the will to will the good. It is therefore a good state of will. But orthodox Christian theology is committed to eschewing Pelagianism.

That is, it is committed to rejecting the view that there is anything good in a human will that is not infused into it by God's grace.

In fact, the second-order desire of the Thomistic will is a part of, or is subsumed under, the volitional component of faith, in which a person longs for God's goodness and repudiates the evil in himself. As traditionally understood, faith is the necessary and sufficient condition for justification. Justification is the beginning of moral and spiritual renewal in a person, and it leads to salvation if only it continues to its end. On traditional Christian views, as long as he continues to have this state of will, a person will eventually come to permanent and everlasting union with God.[53]

Although the rejection of Pelagianism implies that this act of will is infused into a person Paula by God's operative grace, at least from Augustine to Aquinas the Christian tradition has also held that the act of will in question is a free act on the part of the person being justified.[54] Operative grace justifies Paula by producing in her the second-order volition of the Thomistic will, and yet that second-order volition is also supposed to be an act of free will on Paula's part.

The claim that this second-order volition is free and yet also infused in a person by God has been the subject of great controversy. Elsewhere, I have argued that the free will in question can and should be taken in a libertarian sense.[55] In my view, the traditional account of grace and (libertarian) free will can be understood consistently this way.[56]

God is always willing to give operative grace to every person, as long as she does not reject it. If a person Paula rejects it, God cannot give her operative grace without violating her will. On the other hand, since God is always prepared to give this grace, all that is needed for Paula to receive it is her ceasing to resist God. If she ceases resisting and surrenders to God's love, God will produce in her the acceptance of his grace and the second-order volition for a will that wills the good, that is to say, the second-order desire for a will that wills what God wills.

In giving grace to a person Paula, therefore, God is doing all the work, as it were, of producing the relevant second-order will of faith in Paula. But in producing it God is responsive to Paula, who is ultimately responsible for whether or not she has it since it is ultimately up to Paula alone whether she resists God's grace

53 Obviously, on this view, there will have to be some connection between the infusion of God's grace into a person and the passion and death of Christ. There is some contribution to that story in what follows here. For further discussions, see also Stump (forthcoming).

54 See Chapter 13 in Stump (2003).

55 See Stump (2001). See also Chapters 9 and 12 in Stump (2003).

56 For more discussion of this position, see Chapter 12 in Stump (2003).

or ceases resisting it. And so it is right to say that Paula's will is free in a libertarian sense even when the second-order desire of the Thomistic will is produced in Paula by the infusion of God's operative grace. Since it is entirely up to Paula whether she rejects God's grace or ceases rejecting God's grace, it is entirely up to Paula whether or not she has the second-order volition in question.[57]

It is important to be clear that, just as God cannot infuse grace in Paula while she is rejecting it, without violating her free will, so he cannot single-handedly produce in Paula a cessation of the resistance to God. If God single-handedly overrides Paula's will in order to produce a new state of will in Paula, then no matter whether that state is a new second-order will or just a cessation of an old state of will, the resulting condition in Paula is God's will operative in her and not Paula's will. Consequently, if God acts on Paula's will while Paula is resisting God, there will subsequently not be two wills to unite, but only one, namely God's, which is in God but also in Paula. And so the union sought between God and Paula will be precluded, not promoted, if God acts on Paula's will while Paula is resisting God. It is therefore crucial to Paula's justification, and consequently to Paula's salvation, that she herself cease rejection of God and God's grace.

4.2 The Thomistic will and the atonement: Justification

The fact that it is entirely up to Paula whether or not she ceases resisting God's grace does not mean that there is nothing God can to do to help elicit this ceasing in Paula.

The aversiveness of Paula's suffering can be an important part of what moves her towards the Thomistic will and the ceasing to resist grace that is requisite for it. Insofar as God permits the suffering Paula experiences, God plays a role in the eliciting of the Thomistic will in Paula. And no doubt many experiences in a person's life, in addition to suffering, can contribute to readying a person for the transformative experience of surrender to God's grace. Except for cases such as that of Paul on the road to Damascus, perhaps all instances of conversion are preceded by many experiences that soften the resistance of a person to this surrender to God's love. And God's providential ordering of all these circumstances in Paula's life can contribute to softening her heart.

57 This story is compatible with the rejection of Pelagianism insofar as neither the rejection of God's grace nor the bare cessation of rejection counts as a good act of will. The first state is an act of will but not good; the second state does not even count as an *act* of will.

But in order for a person finally to cease resisting God's grace, the heart of a person must not only soften but crack and melt. The notion of a heart's cracking or melting is, of course, a metaphor (just as the notion of softening a heart is). To speak of something's cracking or melting is to describe something's giving way to an external influence after (or in spite of) some internal resistance or disinclination to it. To say that a heart cracks or melts, then, is to imply that a will that previously was resistant or disinclined towards something urged on it by some one (or something) else gives over its opposition and leaves off its resistance.[58]

It is not implausible to suppose that the passion and death of Christ can be a catalyst in bringing a person to this state. If Paula has responded to her suffering by moving in the direction of ceasing to resist God's grace, reflection on the passion and death of Christ can be the catalyst that enables her to go all the rest of the way to abandoning resistance entirely.

Christ's willingness to die for human beings, in their post-Fall condition with all its defects, shows Paula God's great love for her. In the passion and death of Christ, God is not manifesting wrath or rejection of human beings. He is not displaying his regal character, his almighty power, or his role as dreadful judge. On the contrary, in the incarnate Christ God is allowing himself to be put to death in a painful and shaming way because of his love for human beings and his desire to bring them to himself. If anything can help Paula all the way to the cessation of resisting God's grace and surrender to God's love, it does seem that Christ's passion and death could do so. A display of power can prompt fear and submission; but it takes deep love to prompt the kind of surrender to love that precedes operative grace.[59]

For these reasons, one connection between the atonement and the Thomistic will elicited through suffering can be explained in this way. Before a person Paula is justified and God's grace infuses the second-order desire of the Thomistic will into her, Paula has a resistance towards God.[60] When Paula has been readied by

58 I have described the will's state in such a case as if it consisted in ceasing to do an action, rather than as performing the action of ceasing, both because the description of the will as passive seems to me truer to the phenomena and also because Aquinas's philosophical psychology allows for this possibility, which is the basis for the position I am arguing for. For further discussion of this issue, see Chapter 13 in Stump (2003).

59 For the distinction between submission and surrender, see Stump (2010): Chapter 8.

60 In his book Nagel (1997) Thomas Nagel gives a vivid description of the phenomenology of such resistance. Describing the fear of religion, Nagel says,

> "I speak from experience, being strongly subject to this fear myself: I want atheism to be true… It isn't just that I don't believe in God and, naturally, hope that I'm right in my belief.

past experience of suffering (which is itself a result of God's providential working in Paula's life), the passion and death of Christ can be the final means for eliciting Paula's abandonment of resistance. The internal opposition to undergoing the wholesale changes and the humbling entailed by the second-order desire of the Thomistic can give way in the face of the suffering of Christ and the love it shows.

The consequent cascade of events leads to Paula's justification. When Paula's response to the passion and death of Christ results in her surrender to God's love, Paula's giving over resistance to God is followed by God's giving Paula operative grace. In turn, that operative grace infuses into Paula the second-order will in which a person longs for God and God's goodness. This second-order will is in effect a desire to will what God wills. Its presence in a person, together with the accompanying states of intellect, constitutes justification.

In this way, the passion and death of Christ can be instrumental in eliciting from a person the act of will of faith, which subsumes the Thomistic will, and in enabling that person to come to justification, which will lead to union with God if it persists to the end of life.

4.3 The atonement and the gifts and fruits of the Holy Spirit: Sanctification

Once justification has occurred, however, the order of instrumentality is reversed. To see this part of the progression, it is important to understand that the doctrine of the Trinity, which is the hallmark of Christianity, plays a role in the doctrine of the atonement too. On the doctrine of the Trinity, there is only one God but three divine persons, each of whom is the same God while nonetheless being distinct from one another.[61] For my purposes, what is important to understand with regard to this complicated theological doctrine is that, just as there are multiple persons in the Trinity, analogously, if one might put it this way, there are multiple persons in any person Paula who has come to justification.

It's that I hope there is no God! I don't want there to be a God; I don't want the universe to be like that." Nagel (1997), pp. 130–131.

61 The doctrine of the Trinity is, of course, the subject of considerable discussion and controversy, which is not directly relevant to my concerns in this paper. For my purposes, I will just take the doctrine as data for developing the connection between the atonement and the role of suffering in salvation.

That is because once Paula ceases resisting God's grace, God infuses into her not only operative grace but also the third person of the Trinity, the Holy Spirit.[62] The Holy Spirit comes to be within Paula's psyche with the infusion of operative grace and stays there as long as Paula does not return to rejecting God's grace.

In addition, when the Holy Spirit is within Paula's psyche, the Spirit brings into Paula the dispositions and orientations called 'the gifts and fruits of the Holy Spirit'. The gifts are infused dispositions that enhance excellence in Paula's will and intellect.[63] The fruits are the orientations that result from those dispositions and the indwelling Spirit; they include love, joy, peace, patience, long-suffering, and seven others.

A full explanation of the nature of the infusing of the Holy Spirit and the accompanying gifts and fruits is a long story,[64] but the part of that story relevant here is this.

It is part of orthodox Christian doctrine that in his passion and death, Christ somehow took on the sin of all human beings. There are different interpretations of this claim in the history of Christian thought. But one way to interpret it is as a claim that in his passion and death Christ opened himself up to the psyches of all other human beings, all at once, so that he somehow received in himself, in psychic union, the sinful, fragmented, guilty and shamed psyches of all other human beings, without himself actually becoming guilty of any particular sin.[65]

Now union is mutual. In union, one might say, each united person is somehow inside the other with whom he is united. There cannot be union unless the indwelling of love is mutual. So to say that, in his passion and death, Christ opened himself to the sinful psyches of other human beings is in effect to say that, in his passion and death, Christ did his part of what is needed for his union with every human being. His part was to open his psyche up to their indwelling in him, so to say. If Christ had not in this way done what was needed from him for union with human beings, then human beings could not unite with Christ.

But since in his passion and death Christ did his part of the uniting by letting sinful human psyches indwell him, what is missing for union between him and any other human person in this life is just that she allow Christ into her psyche.

62 For a detailed discussion of the notion of the indwelling of the third person of the Trinity in a person in faith, see Stump (2013a).

63 There are three gifts for the will (pietas, fortitude, and fear of the Lord), and four for the intellect (wisdom, understanding, counsel, and knowledge). For a discussion of the gifts and fruits of the Holy Spirit and their place in Aquinas's ethics, see Stump (2011).

64 For a helpful discussion of the gifts and fruits of the Holy Spirit, see Pinsent (2012).

65 For more discussion of this interpretation, see Stump (2012b).

And for that possibility to be realized for Paula, Paula herself has to be open to Christ.

On Trinitarian lore, however, the Holy Spirit and Christ are the same God. There is only one mind and one will in the Holy Spirit and in the second person of the Trinity, for example; and the second person of the Trinity is also the only person in the incarnate Christ. One way to summarize this complicated Trinitarian lore is to say that the spirit of Christ is the Holy Spirit. In his passion and death on the cross, Christ opened himself up to Paula's psyche. In justification, when operative grace is infused into Paula, the spirit of Christ comes into Paula with that grace. So when Paula surrenders to God's love and in justification is open to the indwelling Holy Spirit, the other half of what was needed for mutual union between Paula and Christ is begun.[66]

The gifts of the Holy Spirit that are dispositions for the will then begin the slow work of the sanctification of Paula. That is, they work to bring Paula's will, bit by bit, more in harmony with God's will as regards her first-order willing. They add strength to the part of Paula's first-order will that is in accord with her second-order will to will what God wills. They help Paula to harmony between her first-order and second-order desire, without taking away from her those first-order desires for what she cares about and loses in suffering. And the gifts of the Holy Spirit having to do with the intellect add excellence to Paula's mind, too. They help her mind to be in harmony with God's mind, so that she knows what the will of God is with regard to particular actions and occasions.[67]

More important than the dispositions that are infused with the Holy Spirit, however, is just the fact that Paula has Christ intimately present to her.[68] In the condition of justification, when the Holy Spirit comes to indwell her, Paula has some direct and immediate second-personal experience of Christ and Christ's love for her, however attenuated or tacit the awareness of that experience might be. In fact, the orientations that are the fruits of the Holy Spirit stem from this second-personal experience: love, because Christ who loves Paula is present to her; peace, because in being united with Christ in this way, Paula somehow has her deepest heart's desire stilled; and joy, because Paula's union with the God who loves her is a joyful thing for her. Patience and long-suffering, the next two orientations on

66 For defense of the claim that union between persons comes in degrees, see Stump (2010): Chapter 6.

67 For more discussion of the kind of second-personal knowledge involved in this connection between God's mind and the mind of a human person in faith, see Stump (2014).

68 'Present to' is a complicated notion, hard to spell out just in passing. For a detailed discussion of it, see Chapter 6 of Stump (2010).

the list, arise in consequence. In the presence of a loving God, one can endure suffering, which is still really suffering, without losing love, peace, and joy.

So when the Holy Spirit is infused in Paula, her union with God is begun, not only in the minimal sense that she has a second-order will to will what God wills, but in the much stronger sense that, in the person of the Holy Spirit, God himself is present to her, with her, and, in fact, somehow indwelling in her.

This incipient union is brought about in Paula first because of the passion of Christ on the cross, in which Christ opens himself to union with all human beings, and then because the effects on Paula of her suffering and of Christ's suffering for her elicit from her a giving over of resistance to God's grace. If only Paula does not return to resisting grace, this beginning of union will continue in the process of sanctification. And if Paula does not give up before the process of sanctification is complete, then sanctification will ultimately eventuate in full union between her and God. In the fullness of union, the presence of God to Paula will be greatly more powerful, and Paula's will will be completely integrated around the goodness of God.

The gifts and fruits of the Holy Spirit will help significantly in moving Paula from the incipient uniting with which the process begins to full union with God. But so will the inner, spiritual presence of Christ to Paula. Sanctification is inevitably accompanied by suffering, if in no other way, then because it requires Paula's struggle against the first-order desires that she herself wants not to act on. The indwelling Holy Spirit in Paula, making available to her the loving presence of Christ, will strengthen her for not giving up in the process. She does not have to endure that or any other suffering in her life alone. Christ is Emmanuel for her: God with her. And so she has a powerful help against returning to her former resistance to God.

We can sum up this part of the connection between the atonement and the Thomistic will this way. Once justification has occurred, then the union made possible by Christ's passion and death enable Paula to enter into an incipient union with Christ by means of the Holy Spirit indwelling in Paula. And so, because of Christ's passion and death, the second-order volition of the Thomistic will that is effected in Paula in justification results in the Holy Spirit's indwelling in her. And this, together with the gifts and fruits of the Holy Spirit, helps to bring about Paula's sanctification, which results in her full union with God.

5 Conclusion

On the way of thinking about the connection sketched here between the role of suffering in salvation and the redemptive effects of Christ's passion and death, on the doctrine of the atonement, there is a progression in which the suffering of a person Paula and the passion and death of Christ work together to bring about Paula's salvation.

First, Paula's reaction to her own suffering and her response to the passion and death of Christ can work together to elicit in Paula the cessation of resistance to God's love and grace. That surrender is followed by the formation in Paula of the second-order desire of faith, which subsumes the Thomistic will. Because Christ's passion and death can help to prompt in Paula the giving over of resistance to God, without which God cannot infuse either the second-order will necessary for justification or the Holy Spirit, the passion and death of Christ can be instrumental in bringing a person to justification. In this way, the atonement plays a significant role in synergistically furthering the effects that suffering can provoke in a human person.

But, secondly, once justification has occurred, and God has infused in Paula both operative grace and the Holy Spirit, the process of sanctification starts. The uniting begun by Christ in his passion and death then does its work through the indwelling Holy Spirit together with the infused gifts and fruits of the Holy Spirit. In sanctification, Paula's will remains free, but God increasingly integrates Paula's first-order desires around the good, without destroying Paula's own will or taking away from her what she herself cares about. In consequence, Paula comes more and more to will what God wills; and the incipient union between Paula and God characteristic of Paula when she is first justified increases too.

In this way, the second-order desire of the Thomistic will that suffering helps to elicit is instrumental to making possible for Paula the effects of sanctification provided by the passion and death of Christ.

Furthermore, the work of sanctification would cease if at any point in the process Paula returned to rejecting grace. For sanctification to continue and reach its end of full union between Paula and God, Paula must not revert to rejecting God's grace. That is, she must continue to maintain the second-order desire of the Thomistic will.

So there is a complicated progression that interweaves Paula's suffering and Christ's. Provided only that she does not return to her old habits of resistance, through Paula's suffering and Christ's woven together in Paula's life history, Paula's sanctification will result in her willing what God wills completely, in full union with God. There is a role for Paula's suffering in the progression leading to

her union with God, even though it is Christ's passion and death that makes that salvation available.

In a world in which there is the atonement but no suffering, the chances of a person's availing himself of the salvific benefits of the atonement seem very small.

Every part of this story raises questions and calls out for further explanation. But I hope that this sketch is at least suggestive for a way in which that further explanation might go.[69]

Bibliography

Eckhart, M. (1978a), "Sermon, Blessed are the Poor", in *Meister Eckhart: Mystic and Philosopher*, translated with commentary by R. Schuermann, Bloomington, IN: Indiana University Press, 214–220.

Eckhart, M. (1978b), "Sermon, See what Love", in *Meister Eckhart: Mystic and Philosopher*, translated by R. Schuermann, Bloomington, IN: Indiana University Press, 135–136.

Eckhart, M. (1981), "The Essential Sermons, Commentaries, Treatises, and Defense", in *The Book of Divine Consolation*, translated by E. Colledge and B. McGinn, Mahwah, NJ: Paulist Press, 215–216.

Eckhart, M. (1994), "The Talks of Instruction", in *Selected Writings*, translated by O. Davies, London: Penguin Books, 41–42.

McCord, A. M. (2008), "Plantinga on "Felix Culpa": Analysis and Critique", in *Faith and Philosophy*: 25, 123–139.

Nagel, T. (1997), *The Last Word*, Oxford: Oxford University Press.

Pinsent, A. (2012), *The Second Person Perspective in Aquinas's Ethics*: *Virtues and Gifts*, Routledge.

Plantinga, A. (2004), "Supralapsarianism, or 'O Felix Culpa"', in *Christian Faith and the Problem of Evil*, edited by P. van Inwagen, Grand Rapids, MI: Eerdmans Publishing Co.

Stump, E. (1988), "Sanctification, Hardening of the Heart, and Frankfurt's Concept of Free Will", in *Journal of Philosophy*: 85, 395–420. Reprinted in *Perspectives on Moral Responsibility*, 2003, edited by J. M. Fischer and M. Ravizza, Ithaca, NY: Cornell University Press, 211–234.

Stump, E. (2001), "Augustine on Free Will", in *The Cambridge Companion to Augustine*, edited by N. Kretzmann and E. Stump, Cambridge: Cambridge University Press, 124–147.

Stump, E. (2003), *Aquinas*, London and New York: Routledge.

Stump, E. (2010), *Wandering in Darkness: Narrative and the Problem of Suffering*, Oxford: Oxford University Press.

Stump, E. (2011), "The Non-Aristotelian Character of Aquinas's Ethics: Aquinas on the Passions", in *Faith and Philosophy*: 28, 29–43.

69 I am grateful to John Keller for helpful comments on an earlier draft of this paper.

Stump, E. (2012a), "The Nature of the Atonement", in *Reason, Metaphysics, and Mind: New Essays on the Philosophy of Alvin Plantinga*, edited by K. Clark and M. Rea, Oxford: Oxford University Press, 128–144.
Stump, E. (2012b), "Atonement and the Cry of Dereliction from the Cross", in *European Journal for Philosophy of Religion*: 4, 1–17.
Stump, E. (2013a), "Omnipresence, Indwelling, and the Second-Personal", in *European Journal for Philosophy of Religion*: 5, 63–87.
Stump, E. (2013b), "Conversion, Atonement, and Love", in *Conversion*, edited by I. U. Dalferth and M. Rodgers, Tübingen: Mohr Siebeck.
Stump, E. (2014), "Faith, Wisdom, and the Transmission of Knowledge through Testimony", in *Religious Faith and Intellectual Virtue*, edited by T. O'Connor and L. G. Goins, Oxford: Oxford University Press.
Stump, E. (forthcoming), "Shame, Guilt, and Atonement", in *Essays in Honor of William Alston*, edited by D. Howard-Snyder.
Van Inwagen, P. (1988a), "The Place of Chance in a World Sustained by God", in *Divine and Human Action: Essays in the Metaphysics of Theism*, edited by T. V. Morris, Ithaca: Cornell University Press, 211–235.
Van Inwagen, P. (1988b), "The Magnitude, Duration, and Distribution of Evil: A Theodicy", in *Philosophical Topics*: XVI(2), 161–187.
Van Inwagen, P. (1991), "The Problem of Evil, the Problem of Air, and the Problem of Silence", in *Philosophical Perspectives*: 5, 135–165.
Van Inwagen, P. (1996), "Reflections on the Essays of Draper, Gale, and Russell", in *The Evidential Argument from Evil*, edited by D. Howard-Snyder, Bloomington: Indiana University Press, 219–243.
Van Inwagen, P. (1997), "Probability and Evil", in *The Possibility of Resurrection and Other Essays in Christian Apologetic*, edeted by P. van Inwagen, Boulder: Westview Press, 69–87.
Van Inwagen, P. (2000), "The Argument from Particular Horrendous Evils", in *Proceedings of the American Catholic Philosophical Association*, (*Annual Supplement to The American Catholic Philosophical Quarterly*): 74, 65–80.
Van Inwagen, P. (2004), "The Argument from Evil", in *Christian Faith and the Problem of Evil*, edeted by P. van Inwagen, Grand Rapids: Wm. Eerdmans and Co., 69–87.
Van Inwagen, P. (2005), "The Problem of Evil", in *The Oxford Handbook of Philosophy of Religion*, edeted by W. Wainwright, Oxford: Oxford University Press, 188–219.
Van Inwagen, P. (2006), *The Problem of Evil*, Oxford: Oxford University Press.

Dean Zimmerman

Resisting Rowe's No-Best-World Argument for Atheism

Abstract: William Rowe famously argues that it is impossible to suppose that God exists if, for every world God could create, there is a better one God could have created. Rowe's conclusion suggests an argument for atheism: Since every world could be improved upon, God does not exist. The essay examines two ways to resist the argument: Deny that every creatable world could be improved upon, or deny the normative principle that settling for a less-than-maximally-good world implies moral imperfection.

1 Introduction

One argument – stated succinctly in a one-page article by Stephen Grover, and developed most fully by William Rowe – has generated a great deal of recent work at the intersection of philosophy of religion and value theory.[1] In the first instance, it is an argument that, if there is no best possible world for God to create, then God does not exist. But Rowe's conclusion can be turned into an argument for atheism, since there are reasons to think that an omniscient, omnipotent being would in fact be confronted by an endless series of ever better creative alternatives. I will call this the "no-best-world" argument.

As an argument for atheism, the no-best-world argument does not strike me as very formidable. There are many ways to resist it, each of which will seem plausible to at least some of the philosophical target audience. In exploring these various lines of resistance, philosophers of religion have begun to engage in a good deal of honest toil in value theory and philosophical theology. We have had to pay more attention to what value theorists are saying about commensurability and comparability of values, and about decision when infinite utilities are at stake. As a result, our conceptions of God's goodness, God's creative act, and divine freedom are, I believe, becoming more nuanced.

0 Most of Sections 6–9 originally appeared in Zimmerman (2018); that content is used with the kind permission of the editor of *Faith and Philosophy*.

1 Rowe (1993, 1994, 2002); the fullest presentation is in Rowe (2004), Ch. 6. The core argument appears inGrover (1988).

https://doi.org/10.1515/9783110664812-024

To be sure, deep dialogue with contemporary ethics and value theory does not begin with the responses to this argument. (Especially noteworthy is Robert Adams's essay from 1972, "Must God Create the Best?"[2]) But Rowe's work has certainly drawn more philosophers of religion into this territory.[3]

The chapter constitutes a survey of the two main responses to the no-best-world argument for atheism.[4] I begin, in section 2, with some remarks on the loci of value – the kinds of things that make one world better than another. Next, in section 3, I formulate Rowe's argument, and the argument for atheism based upon its conclusion. That argument rests upon two pillars: (a) the assumption that there is no best world, and (b) the conclusion of Rowe's argument that, if there is no best, there is no God. Sections 4 and 5 consider the plausibility of (a). In discussing possible responses to the assumption that there is no best world, a host of questions arise concerning comparability of values, and the principles for deciding among infinitely good outcomes. The remainder of the paper examines (b), focusing on the second premise of Rowe's argument, and the crucial normative principle supporting it. This principle has been widely criticized; here, I offer some new thought experiments that I take to further undermine Rowe's principle. Much of what I say is not particularly new, though I hope I have a couple of ideas to add to the discussion.[5]

2 Axiological background

Two main strategies have been used to focus attention upon an objective scale of goodness and badness. One emphasizes independence from surroundings, the other emphasizes the worth of pursuing a thing for its own sake. The first leads to the notion of intrinsic value: the goodness or badness of a thing, no matter what setting it is in – whether all by itself, or surrounded with various other things. The second emphasis leads to the notion of final value: goodness or badness "as an end". Some things are worth bringing about, if one can, *only* in order to bring something *else* about; while other things are worth bringing about, if one can, for

2 Reprinted in Adams (1987), pp. 51–64.

3 Recent examples include the contributions to Swenson and Zimmerman (2019).

4 The essay was originally intended to be the introduction to a volume on the problem of evil and infinite value. Many of the essays it was intended to introduce appear (unceremoniously, without introduction) in Swenson and Zimmerman (2019).

5 Langtry (2008) includes an insightful examination of Rowe's argument and its presuppositions (pp. 48–110). My own analysis is indebted to Langtry's, and I largely agree with his conclusions.

their own sake. These are the things that have positive *final* value; they are good in themselves.

These *may* be two names for the same thing.[6] But what matters, for present purposes, is the connection between objective goodness and badness, of one type or another, and what a perfectly good and benevolent deity should be expected to allow or cause. I will remain neutral about whether intrinsic and final goodness and badness impose the very same standard; I assume only that one or the other kind of objective goodness and badness is primarily relevant to what God should do – insofar as God should be moved by ends at all.

What are the bearers of goodness and badness? Is goodness or badness a kind of stuff, like a gas that can fill some rooms and not others? Is evil a sticky substance that gloms onto things, staining them?[7] In his Manichean days, Augustine says, he thought of evil in this way: it is a kind of material or stuff, repelling and repelled by another kind of material or stuff – goodness. The two substances are like light and darkness, and the darkness is trying to drive back the light. The Manicheans could explain its existence: evil is an uncreated thing, battling against good. But it was very hard to see how a God who was as powerful and good as Christians say, and who is responsible for the existence of everything, would create such a substance. Augustine, of course, rejected the Manichean picture of evil-as-substance; he famously replaced it with the view that evil is nothing positive in itself, but always a mere absence of goodness.[8]

Part of Augustine's doctrine that evil is a privation of goodness seems to me to be simply a useful clarification of the metaphysics of goodness and badness. When we say something is bad, we are ascribing a bad *property* to a *substance* or *individual* – but that individual always remains, in another sense, good (at least, according to Augustine). The perfectly valid point here is that we have to keep track of the things we are calling "good" and "bad" – distinguishing between good-making and bad-making properties that might belong to the same thing, in virtue of which the thing can be in some respects good and in others bad. Here is a catalogue of things that can be said to be objectively good and bad, and some of the relations between them.

---

6 Korsgaard (1983) claims they are not. Michael Zimmerman admits the distinction of the two concepts, but argues that they coincide: the final value of a thing is fully determined by its intrinsic value, and only final value is fully determined in this way. See Zimmerman (2001), pp. 60–64.

7 In the movie *Time Bandits*, a satanic character named Evil has been blown up, leaving black bits everywhere. The Supreme Being advises: "Do be careful! Don't lose any of that stuff. That's concentrated evil. One drop of that could turn you all into hermit crabs."

8 For discussion, see Evans (1982), pp. 33–35.

Individuals (*Augustine's Substances*): they can be more or less good or bad, partly good and partly bad; this amounts to being good in some respects and not others.

Properties: Since an individual can be good in some respects (a person may be generous, and kind to children) but bad in other respects (she may also be prone to lie, and disrespectful toward her parents), one must recognize that different aspects or attributes of a person, i.e., different *properties* she has, must also be said to be good and bad; and it is in virtue of having these that an individual is, overall, good or bad.

States of affairs: This is something of a philosophical term-of-art, but it is pretty easy to get the hang of how to use it. Take some individual or group of individuals, actual or merely possible – for example, the students in one of my classes, or a herd of 100 unicorns. Consider some property that they could have but might not have had – for example, taking a test in the same room, or forming a human pyramid. My students' taking a test together, and the same students' forming a human pyramid, are two different states of affairs. The first will almost certainly come about or be the case; the second probably will not come about or be the case. A herd of unicorns' stampeding over a cliff is a state of affairs that could have been the case, but that won't because there are no unicorns. Roughly speaking, a state of affairs is good or bad if its being the case would make the world a better or worse place. Good and bad properties are involved in good and bad states of affairs, respectively. One cannot just look at the properties instantiated, though, to assess the overall goodness and badness of a state of affairs; it may matter which individuals have them – consider, for example, a deserving or an undeserving person experiencing rewards or punishments. Furthermore, overall good states of affairs may involve some things' having some bad properties, and overall bad states of affairs may involve some things' having some good properties. States of affairs that will be of special importance are the biggest, most inclusive states of affairs that are possible: i.e., possible worlds.

Facts or events: A fact or event is just what you call a state of affairs when it is actually the case.[9] One state of affairs that was actually the case is *Caesar's crossing the Rubicon*. Caesar's crossing the Rubicon is a fact; it was also an event that occurred. There isn't much reason to distinguish between these things (facts, events, and states of affairs that actually are the case).[10] When a good state of affairs is the case, it is a good fact or event; when a bad one is the case, it is a bad fact or event. My having a headache at such and such time is a state of affairs that is the case; we can say it is a fact that I have the headache, that my having the headache is an event that occurs. And since being in pain is a bad property, it is a bad state of affairs.

One can argue about whether facts, states of affairs, properties, or individuals should be regarded as "the most fundamental bearers of value". So long as all the intuitively correct connections are retained between the goodness and bad-

9 Thus saith Chisholm, at any rate; see, e.g., Chisholm (1976), Ch. IV.

10 There are complexities here, reasons one might want to distinguish facts from events; but the relationship between event-talk and fact-talk is intimate, and nicely anatomized by Bennett (1988) (see esp. Chs II and VIII).

ness of the entities falling into these various categories, the outcome should not much matter for the purposes of assessing Rowe's argument. There is, however, one point at which it will become important to distinguish among kinds of value-bearer: namely, in identifying the "locations of value" mentioned by "Pareto Principles" (see section 5, below).

The states of affairs that are complete possible worlds should be distinguished from "worlds" in other senses of the word. When states of affairs are conceived as the kinds of things that can exist whether or not they are the case, the "bringing about of a state of affairs" is not a matter of bringing it into existence, but rather a matter of making it to be the case (or causing it to "obtain"). God's choosing a possible world to create should really be conceived of as God's selecting a maximal, consistent state of affairs and making it to be the case. This state of affairs includes everything that is the case, including God's own existence (something that God does not cause, unless self-causation makes sense), and mathematical and logical truths (which Descartes thought God could chose, but his is a minority report); many states of affairs in a possible world are not, then, things that God brings about. Another thing that might be meant by God's creating a world is God's causing a certain sum total of contingent things to exist – all the contingent things there are, according to some possible world. Sometimes it is harmless to slide between these two meanings of "world" when making claims about the range of worlds that God could create. When it matters, however, I shall try to be more careful.

3 Rowe's argument

Rowe's argument relies crucially upon a Principle with the evocative name "B".

> B: If an omniscient being creates a world when there is a better world it could have created, then it is possible that there exist a being morally better than it.[11]

11 Rowe (2004), p. 91.

This is backed up by the intuitive thought that: in general, someone willing to settle for a less good outcome, other things being equal, is morally inferior to someone who has higher standards.[12]

Rowe is careful not to confuse B with B1 or B2:

B1: If an omniscient being creates a world when there is a better world it could create, then it does something morally wrong.

B2: If an omniscient being creates a world when there is a better world it could create, then it wrongs someone.

Rowe does not rely upon B1; he admits that, if the same type of complaint could be lodged no matter what God does, it may not be wrong to do one of those things. He is also not oblivious to a point made by Robert Adams many years ago about obligations to the non-existent, which undermines B2. Suppose some possible world W and set of creatures S are such that: (i) W is the best world that contains all and only these particular creatures, and (ii) each is at least as well off in W as it is in any other possible world. Even if there are better possible worlds than W, with creatures who are better off, merely possible people cannot be wronged; and so, if God were to bring about W, no one would be in a position to complain.[13]

If, for any possible omniscient, omnipotent being, there is an infinite range of better and better worlds for which that being could be responsible; then any God will have to "settle" for one that is not as good as that God could have chosen. According to Rowe's principle B, a being willing to settle for a world with a certain value is not as good as one who is only willing to settle for better worlds. (It is important that any hypothetical creator be forced to choose among a range of better and better worlds; it is not enough for there simply to be no *unique* best world. If there were many tied for best, the argument that follows would not work.)

Here is my reconstruction of Rowe's argument:

1. For any x and W: if x is all-powerful and all-knowing in W, and there is no maximum to the goodness of the worlds available to x, then x could have ensured the existence of a better world than W.
2. For any x and W: if x, in W, could have ensured the existence of a better world than W, then W is a possible world in which x is not as good as x could have been. [Implication of Rowe's Principle B]

12 Rowe (2004), pp. 100–101.

13 Adams (1987), pp. 53–54.

3. So, for any x and W: if x is all-powerful and all-knowing in W, and there is no maximum to the goodness of worlds, then W is a world in which x is not as good as x could have been. [From 1 & 2]
4. There is no possible world W in which God exists but is not as good as God is in some other world W* – nor is there any world in which any other being is better than God is in W. [Moral Unsurpassability]
5. So, for any W: if there is no maximum to the goodness of worlds, then, if God exists in W, God is not all-powerful and all-knowing in W. [From 3 & 4]
6. There is no possible world in which God exists but is not all-powerful and all-knowing. [Essentiality of Omnipotence and Omniscience]

So, for any W: if there is no maximum to the goodness of worlds, then God does not exist in W. [From 5 & 6]

Which is to say, if there is no maximum to the goodness of creatable worlds, God is impossible. Add to this the assumption that, for every world an omniscient being could create, there is a better it could have created, and one has the no-best-world argument: an argument for the conclusion that, for any possible world, God does not exist in that possible world – in other words, atheism is necessarily true.[14]

I begin by raising, and setting to one side, three very general sources of worry about this argument for the necessity of atheism: that it presupposes a crude consequentialism, that it assumes that God's creative act would consist in selecting an entire world, and that Molinism undermines a crucial part of Rowe's reasoning.

Given Rowe's dialectical goals, it would be inadvisable for him to rest his case upon a simple consequentialism, according to which not choosing the best outcome automatically counts as a moral failure. This would certainly decrease the target audience, since there are many non-consequentialists out there, and they probably predominate among those for whom God's existence is a live possibility. (Counterexample: Gareth Matthews once told me, "God is a utilitarian", though I never got round to asking him whether he just meant "utilitarianism is true" or something more complicated – perhaps: God rightly adopts utilitarian moral principles, but we should do something else.)

It seems to me that Rowe can plausibly side-step the charge of assuming a crude utilitarianism. The overall goodness or badness of the consequences of an action must be *relevant* to its moral status. Other things being equal, it should be decisive. Outside of a context that would make the action a lie, or a promise-

14 Rowe himself does not straightforwardly endorse the no-best-world argument; but only the dilemma that either God does not exist, or God "cannot enjoy much in the way of libertarian freedom with respect to creation" (Rowe (2004), p. 7). Nevertheless, I shall sometimes sloppily refer to the entire no-best-world argument as "Rowe's argument".

breaking, or … (i.e., so long as all so-called "side-constraints" are satisfied), the value of the outcome should determine its choice-worthiness. And it is reasonable to suppose that better rational beings should be attracted to better outcomes.

A second worry is that the argument targets only a subset of today's theists – namely, those who believe God's creative act is rightly regarded as the selection of an entire world, all at once. Calvinists and Molinists believe that God can choose a whole world (with free creatures in it), and ensure that it comes about; but others deny that God's creative act includes selection of the world in all its detail. If God does not determine everything, and Molinism is not true, then it is not at all clear that what an omniscient and omnipotent being does is "choose a world", at least not if free creatures are involved or any events are left to chance.

The most popular alternative accounts of providence and foreknowledge (alternatives to the accounts available to Calvinists and Molinists) are: Simple Foreknowledge, Timeless Choice, and Open Theism. To simplify the description and comparison of these views, I shall assume that the created world has a beginning. According to advocates of Simple Foreknowledge and Timeless Choice: God chooses certain initial conditions for creation, and also makes conditional choices (or, better, choices about conditionals) – that is, God decides what will happen if so-and-so freely does such-and-such, for all the alternatives that will or could be left open to creatures. The result is a complete world which God has either always known about in full detail (Simple Foreknowledge) or timelessly knows about in full detail (Timeless Choice). Open Theism limits foreknowledge: God does not know genuinely chancy outcomes until they happen, but could still set up initial conditions, and conditional choices about what to do if … . (God could also "play it by ear".) In any case, on these three views, the states of affairs that God can unilaterally ensure will be the case are not complete possible worlds.[15]

Do these views interfere with Rowe's reasoning in the no-best-world argument? After all, on the latter views (Open Theism, Simple Foreknowledge, and Timeless Choice), there might be better possible worlds than the actual one, but their failure to be chosen would not be God's fault.

In the next section, I briefly survey typical reasons for thinking there are better and better worlds, with no best, under the assumption that what God does is to choose an entire world. If these are good reasons, I suspect that they could be reformulated so as to apply to the alternatives on which God does not choose an

15 For comparison of Open Theism, Simple Foreknowledge, Molinism, and Calvinism, see Beilby and Eddy (2001). What I call "Timeless Choice" would involve a theory of divine knowledge in which there is more than one "stage" (compare the theory of stages developed in Zimmerman (2012)).

entire world. So I shall put this issue to one side as well. (Though I will worry about it again, briefly, below.)

As Brian Leftow has pointed out, if Molinism is true, there might well be a best possible world, though it may not be accessible to God.[16] In those circumstances, it would simply be bad luck that God cannot get to it, and so not choosing it would be no reflection on God's character. Furthermore, if the Molinist is right, there might well be an infinite hierarchy of better and better worlds, but a limit to the goodness of the worlds God can actually choose. What this shows is that, in the no-best-world argument, talk about worlds "available" to the being in question should be taken to be those that are (in Flint's terms[17]) "feasible", something that – according to the Molinist – varies from world to world. Whether the set of worlds available to a being has a maximum is, at least potentially, a contingent question and must be relativized to a world throughout the argument.

One way to modify the argument so as to take into account the possibility that the absence of a maximum available world might be world-relative would be to start with the premise:

> 1*. For any x and W: if x is all-powerful and all-knowing in W, and there is no maximum to the goodness of the worlds available to x in W, then x could have ensured the existence of a better world than W.

Carrying the change through the rest of the argument, one would arrive at the conclusion that: "for any W: if there is no maximum to the goodness of worlds available in W, then God does not exist in W." Although this does not imply that, if there is no best possible world then there is no God, it does imply that if there is no best possible world then God does not exist necessarily. Those who want to maintain God's absolute necessity will need to resist this version of the argument.

Another way to respond to the worry that Molinism derails the argument would be to argue that, however the Molinistic conditionals turn out, there would always be an infinite hierarchy of better and better worlds available to God – so no modification of the argument is necessary. The question deserves more consideration: do the reasons to be adduced for thinking there is no best possible world also support the conclusion that, no matter how the Molinistic conditionals turn out, there will be no best feasible world? I shall not take up that question here.

16 Leftow (2005b).

17 Flint (1998), pp. 51–54.

4 Simple reasons to think there is no maximally good world

If you could always add some more good stuff, without detracting from the value of what you already have, then there will be better and better worlds with no best. If there are no "organic unities"[18], and suitably individuated valuable states of affairs can always be aggregated, then the antecedent seems pretty obvious, and one reaches the conclusion that there is no maximum to the goodness of worlds available to God.

Suppose there are organic unities, so that you can add 10 points of goodness and not thereby add 10 points to the total value of a larger state of affairs. Still, it might the case that, for every possible world, there are some kinds of good stuff that can be added to generate a different, larger possible world that is *somewhat* better, even if there are diminishing returns. Adding spatiotemporally disconnected valuable things might well have this effect. In that case, the possible worlds amongst which God must choose form a hierarchy that becomes better and better, ad infinitum. And that is good enough for the no-best-world argument.

When aggregation of value fails, and whole is less than the sum of the parts, it must be due to the fact that there is something improper or inappropriate about the arrangement of the value bearers. So long as there is nothing inappropriate or bad about disconnected universes, or large manifolds of arbitrarily large sizes containing isolated universes, or a universe going on and on forever with ever more good stuff; then adding more and more to a universe of finite value ought also to lead to good worlds of larger and larger values. If no worlds had infinite value, then this line of thought would support the conclusion that the hierarchy of possible worlds has no top, even if organic unities are responsible for diminishing returns.

5 Complications

But worlds might have infinite values; and that leads to serious complications, and a variety of reasons to doubt that God would be confronted with a never-ending choice of better and better worlds to create. In this section, I survey – inconclusively and incompletely – some of those reasons.

18 For defense of the existence of what G. E. Moore called "organic unities", see Lemos (2015).

Any world God could create would include the state of affairs of God's existing. Do we want to say any world with God in it is infinitely good? Suppose we do, and this implies that the state of affairs of God's existing has infinite value. Suppose also that higher infinities of value are not possible. If so, and if all worlds are comparable in value, forming a hierarchy; then all the worlds with God in them (and any other worlds that are infinitely good, if such there be) will be at the top, of equal value. Only worlds without God (if such there be) will be lower down. Then there is just one big set of best possible worlds available to God, differing with respect to the amounts of good in the created parts, but equally good overall.

Even given all these assumptions, it ought to be possible to recover Rowe's question, and reinstitute the argument, by subtracting the value that accrues to a world simply from God's existence in that world. When worrying about comparisons of worlds with infinitely many "locations of value", under the assumption of additivity, Vallentyne and Kagan (and others) have defended what are called "Pareto Principles".[19] Working out the details of such principles is difficult, to say the least; and some despair that it can be done.[20] But it is not implausible to think that one should be able to compare at least some worlds with respect to goodness and badness, despite the presence of infinite values. At least sometimes, such comparisons should be possible – for instance, when all the same bearers of value exist in both worlds, and all are at least as good or better in one than they are in the other. Suppose persons are the basic locations of value. When God and a bunch of created persons exist in one world, and in another God and the same individuals exist, all better off than they were in the other, the latter world certainly seems like it ought to be judged better than the first.

Problems about comparing infinitely good worlds will arise shortly, so I postpone further discussion until then; I note here only that, to the extent that Pareto Principles are plausible, they should allow for the bracketing of God's purportedly infinite value in the comparison of worlds in at least some cases.

But what if some of the goods that make a world better are not comparable, or not commensurable? In that case, how should one think about the question whether there is no maximum to the goodness of the worlds available to God? Suppose, for example, that either aesthetically valuable states of affairs are never better or worse than dutiful actions; or that, although some comparison can be made – say, every dutiful action is better than any aesthetically valuable state of affairs (or vice versa, according to the truly decadent) – no cardinal unit of com-

19 See, e.g., Vallentyne and Kagan (1997).

20 For one criticism, see Hamkins and Montero (2000); for a response, see Lauwers and Vallentyne (2004).

parison exists. If some goods are incomparable, there will be worlds that cannot be compared with respect to overall goodness. If some goods are, though perhaps comparable, nevertheless incommensurable, then worlds with both kinds of goods will not have an overall value that can be worked out from adding up their individual values. If worlds nevertheless have a complete ordinal ranking in terms of better and worse, it will have to come from somewhere else.

If consequentialism is false, there might be, even in the case of incomparable worlds, a complete ordering from the point of view of *what a creator should do* – an ordering in terms of choice-worthiness. If the factors that make for greater choice-worthiness admit of arbitrary augmentation, God would once again be in a no-best-world situation.

But suppose choice-worthiness is a function of goodness; and there is no single ordering, because not every world is comparable. Some hierarchies of worlds – maximal comparable collections – might have a best; others not. Bruce Langtry calls the best in a hierarchy, when there are several hierarchies, a "prime" world. He argues: if confronted with one hierarchy with no best, and another hierarchy with a prime world, God should pick the prime world.[21] But why think that? Robert Adams would be unlikely to see the virtue in choosing the prime world, rather than some world from the other hierarchy. Choosing from the no-best-world hierarchy might afford God the opportunity to exercise grace, in Adams's use of the term – love for creatures not based upon degree of dessert.[22] And that could serve as a reason to pick a good enough world from the hierarchy with no prime world, rather than the prime world.

Whether or not God should pick a prime world, given the option, there might be no prime worlds because no hierarchy has a top. The earlier considerations in favor of no best world – that, for the most important goods, one could always add more – suggest infinities will always arise. The best cases against this come from *aesthetic* values; but goods of other kinds arguably trump those, so hierarchies that include non-aesthetic values seem likely to have no top.

This possibility reveals an alternative to resisting the no-best-world argument by positing exactly one best possible world, or a bunch of maximally good equally valuable worlds. One could instead suppose that there are many prime but incom-

21 Langtry (2008), p. 108.

22 Indeed, Adams thinks God could even graciously choose a lesser world from a prime world's hierarchy; "there is nothing in God's nature or character which would require him to act on the principle of choosing the best possible creatures to be the object of his creative powers" (Adams (1987), p. 57). The idea that God could love creatures in a way that provides a reason for creating them is defended in Leftow (forthcoming).

parable worlds amongst which God could choose.[23] However, if each of the hierarchies – each family of comparable worlds – has no best, then Rowe's argument kicks in once again. (From now on, when I ask whether a series of worlds has a best, or its goodness approaches a limit, etc., assume that the series constitutes a hierarchy.)

If morally valuable states of affairs can never have their value trumped by aesthetic considerations such as elegance or clutter, then adding morally good states of affairs – either by adding spatial or temporal extent to a single universe, or to a manifold of universes, disconnected or connected – will always give a world a plus. This assumption does not immediately lead to the possibility of worlds with infinite value due entirely to the created goods within them. Some will say that actual infinities of individuals are impossible – but that is something I reject. So the possibility of infinitely many persons involved in morally good states of affairs represents yet another challenge to the assumption that worlds can always be improved upon: Set God's value aside, and just focus on contingent locations of value; assume that adding goods is always possible, that there are no such things as organic unities, and that infinitely many locations of equal value can be created. This would seem to lead to the existence of many possible worlds for God to create, all infinitely valuable, all equally good. Could the no-best-world argument have any purchase, in this setting?

Once one allows for real infinities of contingent locations of value, it is hard to resist the idea that infinitely valuable states of affairs could be brought about by God. To suppose that the blessed, living forever, would not generate infinitely many locations of value would require that each day of blessed existence has positive value up until a certain point, but then their going on forever is not ideal – that value fades away or cuts out or something. Yet eternal life is supposed to be awesome (and it *would* or at least *could* be – pessimism about the value of eternal life seems to me to be mostly sour grapes). Or take the case of infinitely many happy angels; there is no reason to suppose that their happiness is not additive, and that their existence would not be infinitely good. Many have thought that a maximally perfect being would settle for nothing less than infinitely many valuable things in the created world; and that does seem quite plausible.[24]

Suppose there are infinitely many locations of value with finite worth; and any finite number of them can be added up, so that the overall value can only be infinite. What does this do to the no-best-world argument? If all the infinitely good worlds are equal in value (there are no higher infinities when it comes to

23 See Pruss (2017) for a good defense of this option.

24 For a recent argument for this venerable conclusion, see Climenhaga (forthcoming).

goodness), and all preferable to any of the finitely good worlds, then it looks as though God is off the hook; there is a maximum value for creatable worlds, and God just has to choose one of those.

Admitting worlds with infinitely many locations of value introduces lots of puzzles. Infinitely many goods and evils with finite worth can appear in worlds where the idea of adding them up is problematic. A little trial, -1, is necessary to an overall good of courage or fortitude or something, +10; the result, if not an organic unity, is +9 for one person. But take an infinite series of such people, and try to add up their collective trials and virtues to reach their overall value. The result might at first seem to be simple: -1+10-1+10-1+10.... The partial sums of the series, in that order, are: 9, 8, 18, 17, 27,..., so taking their limit as the result of adding them together yields positive infinity. But that is not the only sum of those numbers. They can also be arranged in this order: 10 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 + 10 - 1 - 1 ..., a series that has no sum; it is not well defined. And they can be rearranged as 10 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 + 10 - 1 - 1 ..., which has negative infinity as its limit. But the world described is surely as creatable as one where a slightly less valuable sort of +9 angel appears infinitely many times – a collection of goods which does have a sum. And one wants to say that the infinity of angels should be as good as the infinity of people with their trials and virtues.

Indeed, we tend to have pretty strong reactions about the comparative values of *some* worlds with infinitely many locations of value, some much stronger than this comparison. If we trust these judgments, we may be able to regain the result that there are better and better worlds with no best, even though all the worlds in question are infinitely valuable.

Peter Vallentyne and Shelly Kagan and are convinced that some such comparisons must be possible. They begin with the following "Basic Idea" for making comparisons among certain worlds – ones with the same locations of value:

> BI (basic idea): if wl and w2 have exactly the same locations, and if, relative to any finite set of locations, wl is better than w2, then wl is better than w2.[25]

There is now an industry of refining and criticizing "Pareto Principles" like this one – principles that allow for the ordering of worlds with infinitely many locations of value.

One question about the interpretation of Pareto Principles is the nature of the "locations of value". Kagan and Vallentyne present several options among which they do not choose: "Locations for local goodness might be points or regions in

25 Vallentyne and Kagan (1997), p. 9.

time, space, or space-time; or they might be people or states of nature."[26] The question whether all the locations of value in one world are identical to those in another becomes very important, because the way this question is answered will result in very different judgments about which worlds are comparable according to these sorts of Pareto Principles. Returning to the list of potential fundamental bearers of intrinsic value in section 2, I note that properties, states of affairs, and facts do not seem likely candidates for locations of value – at least not if these entities have the essential properties they are alleged to have by most metaphysicians. The property of having a very precise sort of pain is not – and could not be – identical to the property of having a stronger or weaker pain. And, at least as states of affairs are ordinarily individuated, the state of affairs of a particular person feeling one sort of pain is not the same as the state of affairs of that person feeling a stronger or weaker pain. For Pareto Principles to generate interesting results, what one needs are locations that can exist in more than one world while varying along the dimensions relevant to their goodness and badness. I shall assume that it is individual substances – persons, animals, plants, rocks, etc. – that are candidate locations of value.[27] They might well not be the fundamental bearers of value – that is, they may not be the things at the foundations of the true axiological theory, the things the goodness and badness of which serves to explain the goodness and badness of everything else. That might be the job of states of affairs, for example. However, the kinds of bearers of value relevant to Pareto Principles might not be the most fundamental bearers.

Nick Bostrom has raised serious objections to all refinements of Vallentyne and Kagan's Basic Idea, especially when they are thought of as useful extensions of summation for consequentialists. For one thing, even where Pareto Principles seem to provide a plausible ranking of situations with infinite positive values, they do not generate anything more than an ordinal ranking; and when making decisions under uncertainty, you need more than that so you can work out expected utilities.[28] If God chooses entire worlds with certainty, as on Calvinism or Molinism, this will not be relevant. Even if there is not a well-defined utility function by means of which to compare two worlds, the rule "choose the better if you can" remains plausible. On the chancy views of providence, what God chooses among are world-types-with-divine-plans-of-response; and here Bostrom's point becomes relevant. God's choice is made with knowledge of the odds of getting this world or that world from a given world-type-cum-plan; but, on such views, the ab-

26 Vallentyne and Kagan (1997), p. 5.

27 For criticism of the use of times, see, e.g., Climenhaga (forthcoming), pp. 14–18.

28 Bostrom (2011), section 2.2.

sence of a utility function would prevent comparison of these choices in terms of their overall utility – one cannot look at a given choice, multiply the values of the goods of the various possible outcomes by their odds, and add the results in order to compare the choice with others. Still, even here some comparisons certainly seem to make sense – for example, when all the possible outcomes of one choice are better than all the possible outcomes of another.

Pareto reasoning can generate a ranking of at least some series of worlds, they could bring the no-best-world argument back – so long as, from any plausibly creatable world, there arises a series of better and better worlds God could choose.

The least controversial Pareto Principles apply to worlds with the same locations of value. Suppose those are the only worlds where direct comparisons can be made. If there is no comparing worlds with infinitely many positive value locations but *different* locations, then the only worlds that will generate a no-best-worlds hierarchy are ones in which there is no maximal amount of good that can be given to the very same individuals one finds in that world.

This suggests another way to resist the no-best-world argument: Admit that creatable worlds are really only ones with infinitely many locations of value; and insist that ranking of worlds with infinitely many locations of value is only genuinely possible when comparing worlds with the same locations in them. And for every location (whether these are persons, space-time regions, or something else), there is only so much value that can be given to it. God has chosen one co-possible set of locations, and conferred upon each of them as much value as it can bear. (That is a big assumption! And it seems hard to believe … but maybe sufficient thought about locations of value might make it a little more plausible – for example, if locations are entire lives, and our lives are much longer than they appear to be, each might be, in the end, as good as it could be…somehow.[29]) But, on this hypothesis, God *could* have created strictly more good stuff, and it is no skin off divine goodness that God did not do so, because those larger worlds are not comparable to this one.

This would make it possible to respond to the no-best-world argument by denying that God could have created a better world: God created a world with *these* goods, when God could have created one with more; but that is not evidence of God's willingness to settle for less, since the worlds are not strictly comparable.[30] There are many prime worlds available to an omnipotent, omniscient

29 For extended defense of a thesis in this vicinity, see Climenhaga (forthcoming).

30 Climenhaga (forthcoming), though aimed at a slightly different atheistic argument, provides the materials for a response to Rowe along these lines (which Climenhaga tentatively endorses; see pp. 23–24).

being; and they are all the creatable worlds, each infinitely good and strictly incomparable with the others. There are many "maxes"; but, since they are not comparable, there is neither a maximum degree of goodness, nor an endless hierarchy of better and better.[31]

One should like to be able to extend Pareto reasoning to at least some cases in which comparable worlds with infinitely many locations of value do not share all the same locations.[32] In the finite case, worlds exactly the same except for the presence of some extra happy people are bound to be better. If any comparisons are possible between worlds with infinitely many locations of value – and it seems plausible to me that some are comparable – cases of the mere addition of some extra happy people seem ripe for comparison. If so, the plausibility of some incomparability among worlds with distinct individuals, in the infinite case, will not be relevant to whether all worlds can be improved.[33] In that case, the no-best-world argument once again would seem to gain a foothold. Suppose that, if God creates contingent things at all, God creates a world with infinitely many locations of value. Suppose further that, for, every world W with infinitely many locations of value, there is a W* with all the same locations, each at least as well off, *and a few more* with positive values. It is extremely tempting to say that, in that case, if God creates anything, God could have created something better. One might well doubt whether these suppositions are true, and whether Pareto reasoning can be extended in this way.[34] But I find the suppositions fairly plausible, and the comparative judgment intuitive.

So, where are we? There is obviously much more to say about all these complications; and many philosophers are addressing the issues. I have just scratched the surface of the extensive literature on Pareto Principles, some of which is highly relevant; and I have not discussed Kraay's argument that there *is* a best possible world, namely the theistic multiverse – the possible world in which God creates every universe above a certain threshold of goodness.[35] I conclude this section simply by reporting that, so far, I have not found conclusive reason to deny that, if God creates any contingent things, there will be a possible world in which what

31 Pruss (2017) argues that there are many sources of incommensurability among worlds, making it plausible that there are many prime worlds.

32 See Vallentyne and Kagan (1997), p. 20, note 18, for an early statement of optimism about such extensions.

33 Philip Swenson agrees, and explores a wide range of possible moral principles that might be thought to govern a deity in a no-best-worlds situation, in Swenson (2018).

34 See Climenhaga (forthcoming), pp. 8–14 for objections to such extensions in the present context.

35 Kraay (2010); but see the criticisms of Monton (2010).

God creates is better. That is reason enough to scrutinize the second premise of Rowe's argument, and principle B which lies behind it.

6 Doubts about Principle B

Principle B is a kind of "expression principle"[36] – it links a psychological state or character trait that comes in degrees, on the one hand, with its expression in behavior under idealized circumstances, on the other hand. Most critics of Rowe's argument respond to it by claiming that, when faced with an endless sequence of better and better options, choosing a worse world than one could is compatible with moral unsurpassability. When circumstances present ever better options from which to choose (what Chris Tucker calls an "Ever Better Situation"[37]), the value of one's choice will not precisely express the goodness of one's motives or character.[38]

This criticism of B will seem intuitive to those already convinced, by similar examples, of the appropriateness of "satisficing" – choosing a good enough option when one could do better. Michael Slote famously argues for a "Satisficing Consequentialism" partly on this basis. He first points out that *rationality* seems not to be impugned by choosing a good enough option from an endless series of better and better options:

> How do we react to fairy tales in which the hero or heroine, offered a single wish, asks for a pot of gold, for a million (1900) dollars, or simply, for (enough money to enable) his family and himself to be comfortably well off for the rest of their lives. In each case the person asks for less than he might have asked for, but we are not typically struck by the thought that he was irrational to ask for less than he could have … .[39]

36 Compare Morris's use of the term "Expression Thesis" (Morris (1993), p. 242).

37 Tucker (2016), p. 130.

38 Representatives include: Morris (1993), pp. 243–245; Howard-Snyder (1994, 1996); Hasker (2004), pp. 171–173; and Langtry (2008), Ch. 2. Norman Kretzmann foresaw Rowe's argument, and made the same move (Kretzmann (1990), p. 238). A different sort of challenge to principle B has been developed by Brian Leftow. He argues that even God might be subject to moral luck, and wind up with a less good world for reasons that do not impugn God's character (I say "largely", because the argument includes several dilemmas, and not all the horns ascribe moral luck to God); cf. Leftow (2005a,b).

39 Slote (1984), p. 147.

Slote then draws a similar conclusion about the *moral* status of someone forced to choose among options that benefit others to higher and higher degrees, without limit:

> [C]onsider a fairy-tale wish regarding people other than oneself. A warrior has fought meritoriously and died in a good cause, and the gods wish to grant him a single wish for those he leaves behind, before he enters Paradise and ceases to be concerned with his previous life. Presented with such an opportunity, may not the warrior wish for *his family* to be comfortably well off forever after? And will we from a common-sense standpoint consider him to have acted wrongly or non-benevolently towards his family because he (presumably knowingly) rejected an expectably better lot for them in favor of what was simply good enough? Surely not.[40]

As Chris Tucker has pointed out, Slote's reaction is widely shared; satisficing in Ever Better Situations has seemed highly intuitive to many moral philosophers – so intuitive as to be "banal", and hardly worthy of the name "satisficing".[41] According to Tucker, it should rather be thought of as a kind of "motivated submaximization" – a situation in which, while aiming at as much good as possible, one settles for less than one could because of "countervailing considerations".[42] Tucker argues that theists who reject B are merely committed "to the appropriateness of motivated submaximization", which comes at little cost since it is "so popular and well supported in the mainstream literature".[43]

Peter van Inwagen has argued that, if God wants a world with certain human goods, satisficing will be acceptable because of the vagueness of what is good and bad for human beings. Van Inwagen notes that, when it is vague whether a certain number of days in prison has the power to deter a certain population from committing a crime (what exactly counts as "deterrence"? when are potential offenders "sufficiently afraid"?), a good judge is forced to draw an arbitrary line – thereby imposing a sentence that is more harsh than strictly necessary. If the worlds among which God chooses differ in goodness in a way that requires a similar drawing of an arbitrary line, then Rowe's Principle B is false.[44]

40 Slote (1984), pp. 150–151.

41 Tucker (2016), p. 130.

42 Tucker (2016), p. 130.

43 Tucker (2016), p. 138.

44 And van Inwagen tells a story intended to show that at least some of the kinds of horrors one finds in our world could show up in a world in which it is vague how many are necessary for God's good purposes. See Van Inwagen (2006), pp. 99–111.

I agree with the critics of B and will attempt to bolster confidence in its rejection not by directly engaging with its defenders,[45] but by telling a couple of stories. The stories raise problems for analogous expression principles that might be thought to govern other personal characteristics besides moral goodness, benevolence, or purity of motivation. These morally charged character traits come in degrees because they depend upon psychological dispositions that also come in degrees – e.g., some people are more willing to make personal sacrifices for the sake of others in need, and qualify as more benevolent in virtue of that fact. The hypothesis that moral perfection is a coherent notion depends upon the idea that there is at least one combination of these psychological dispositions that makes for maximal moral goodness. But should we expect these kinds of gradable psychological traits to be capable of full expression, even in ideal circumstances, when there are infinitely many better and better options available to an agent, with no best?

Reflection upon a couple of other gradable psychological states strongly suggests a negative answer. Certainty and desire come in degrees, and they may well have upper limits – at least, there may be highest degrees of certainty or desire available to creatures like us. We naturally expect that, in ideal circumstances, different degrees of certainty and desire would be expressed in different kinds of behavior. But when agents are supposed to be capable of contemplating infinitely many acts, or even indefinitely many acts, B-like principles for the expression of certainty and desire lead to analogues of Rowe's argument. The results are highly counterintuitive, and this supports the general claim that simple expression principles, like B, should be expected to fail in cases relevantly like an Ever Better Situation.

7 Expression principle for certainty

I begin with the failure of an expression principle for certainty – a principle that seems very plausible, until one considers the infinite case.

In general, the amount of money one is willing to bet reflects one's certainty about the outcome. If an ideal gambler is in ideal circumstances – she is clear-headed, not making any mathematical or logical mistakes, acting entirely out of self-interest – her willingness to risk larger and larger sums should be expected

45 E.g., Kraay (2006), which defends B against the criticisms of the Howard-Snyders.

to go up with her degree of certainty. So this expression principle about idealized gamblers seems, prima facie, reasonable:

B*: If an ideal gambler in ideal circumstances bets *n* dollars on a given outcome, but could have bet more than *n* dollars, then it is possible for the gambler to be more certain about that outcome.

Suppose there is a cosmic lender, Donna, who has convinced an ideal gambler, John, that, for any sum of money he can think of, she will give him that much money to bet on the outcome of a game. Donna plays the role of "The Lender", a role that requires convincing any reasonable person that you have any amount of money the person can name.

Now suppose the game is this: there is a mathematically challenged rich man, Donald, whom the gambler believes to be as rich as Donna; and Donald bets John that 2+2=5. John is *absolutely certain*, as certain as it is possible for John to be of anything, that Donald is wrong about this. John has several thousand dollars to bet, but – because the situation is an ideal one, in which John is purely self-interested, and there are no other relevant considerations – John would bet anything he can get his hands on that Donald is wrong. However, Donna, the Lender, is nowhere to be found. So John bets Donald $1000 that 2+2 is not 5. Donald says, "That's not very much. Let's make this bet interesting; for any amount of money you can name, I'll bet you that much money that I'm right and you're wrong." Suddenly, Donna enters the room, and offers to cover John's bet. John, relieved, chooses the first astronomical amount he can think of – three-hundred-billion dollars – although John could easily have chosen a larger number. While selecting this number, John does not lose any confidence in his knowledge of mathematical facts; he knows he must choose some amount, and that his choice will have to be to some degree arbitrary – not correlated precisely with his rational expectations. John wins the bet, and becomes fabulously wealthy.

If B* is true, there is a deep incoherence in this story, one that can be revealed by an argument that mirrors Rowe's:

1*. For any x and W: if x is an ideal gambler in W and x borrows *n* dollars from the Lender to bet on an outcome, then x could have bet more on that outcome – for example, x could have asked for *n*+1 dollars.

2*. For any x and W: if x, in W, is an ideal gambler who could have bet more on an outcome, then W is a possible world in which x is not as certain of the outcome as x could have been. [B*]

3*. So, for any x and W: if x is an ideal gambler in W and x borrows n dollars from the Lender to bet on an outcome; then W is a world in which x is not as certain of the outcome as x could have been. [1* 2*]

4*. In W* (the world of this story), John is an ideal gambler who is as certain of the outcome as he could be of anything – there is no possible world in which John is more certain about something.
5*. So, for any W: if W is a world in which John borrows n dollars from the Lender to bet on a certain outcome, then John is not absolutely certain that the outcome will occur (John is not as certain as he is in W*). [From 3* 4*]
6*. But W* is a world in which John borrows n dollars from the Lender to bet on a certain outcome (that is part of the story).

CONTRADICTION: In W*, John is not as certain about the outcome as John is in W*.

What this argument shows is that, if B* is true, the conclusion of the story about John, Donald, and the Lender was deeply incoherent. Once Donna offered to cover John's bet, the story could only be continued in one way, so long as John remains an ideal gambler in ideal circumstances (John does not become confused or distracted, John does not cease to be purely self-interested): "John's absolute certainty about the simplest mathematical facts faded away in the presence of the Lender." But there seems nothing inconsistent in finishing the story with John's making an arbitrary choice, and denying that John's certainty changes.

Slote's first fairy-tale moral seems exactly right: When it is better to choose arbitrarily among excellent options rather than not to choose at all, the rational person will choose arbitrarily. There might be some amounts that it would be positively crazy for John to choose, and some that are borderline; but as long as there are many that are perfectly excellent, choosing one when a better is available is consistent with John's being an ideal gambler – he is not making any mathematical or logical errors, he is driven only by self-interest.[46] And the kind of information gained when Donna enters the room, which has nothing to do with whether two and two make five, should have no effect upon John's level of certainty.

46 Tucker cites several affirmations of the rationality of this sort of "first pass satisficing" (i.e., the rationality, in Ever Better Situations, of choosing an arbitrary but highly prized outcome), including: Pollock (1983), pp. 417–418; Schmidtz (2004), pp. 41–44; and Slote (1989), pp. 110–123 (Tucker (2016), p. 130). I would add that even James Dreier, who explicitly argues against the idea that the norms of rationality permit satisficing (Dreier (2004)), would allow that John's choosing an arbitrary amount to bet is more rational than John's not betting at all. Dreier is inclined to say that John faces a rational dilemma: whatever John bets, he can be faulted from the standpoint of practical reason; he is not *perfectly* rational (personal communication). As Daniel Rubio has pointed out to me, this conclusion does not save an expression principle such as (B*), which is not about whether satisficing in Ever Better Situations is compatible with being certain and being perfect rational, but rather whether it is compatible with being certain and being entirely self-interested and logically infallible.

Surely the better thing to say is that 2* and its justification, B*, the certainty analogue of B, should not be expected to hold when a gambler has infinitely (really, in this case, indefinitely) many options from which to choose, and no best option. When confronted by infinitely many (or indefinitely many) choices of a sort that *normally* can be counted upon to precisely reflect or express an agent's level of certainty, one should no longer expect the choice to express the agent's certainty.

8 Expression principle for desire

Here is a similar case in which a different sort of mental state – degree of desire to reach some place – might be expected to be expressed by a different kind of action – the speed with which one travels. On the face of it, an analogue expression principle is plausible (once the case is set up just right): degree of desire should, other things being equal, be expressed in the speed with which one satisfies the desire. But, in the infinite or indefinite case, one should no longer expect the relevant expression principle to hold.

Suppose that an ideal traveler – someone who only cares about the time it takes to get to her destination, and who makes no logical mistakes in her practical reasoning – wants to reach New York quickly, and wants this as much as she could want anything. Since she wants nothing else that could conflict with this desire, she will use the fastest method, and spare no expense. Someone who claims to be in this situation, but who takes the Queen Elizabeth or makes a stopover in Boston to see relatives, does not after all have this maximal desire to get to New York quickly, or else she is not an ideal traveler. This suggests an expression principle:

> B**: If an ideal traveler is willing to settle for a trip to her goal at a certain number of meters-per-second, but could have chosen a faster speed, then it is possible for her to have a stronger desire to reach her goal.

Now take someone who really does have this extreme desire, someone who desires to reach a loved one in New York City as much as she could want anything; and convince her that the following is true:

> Our scientists have created a teletransporter that cannot get you to New York at the speed of light, but it can get you there at any speed you choose short of that. The speed of light in a vacuum is 299,792,458 meters per second. You can choose 299,792,457 meters and 9 decimeters per second, or 299,792,457 meters and 99 centimeters, or 299,792,457 meters and 999 millimeters or … . Now tell us how fast you want to get there.

Of course normal people will think that such tiny differences in speed could never matter to them or their loved ones, since the differences could never be noticed by themselves or their friends in New York. But if our idealized agent believes, rightly or wrongly, that any fraction of a second could be detected by the person she is trying to reach, and that every fraction of a second apart is a sad one for that person, then "it doesn't matter" is not an option. For any ".999999 ..." chosen, something faster could have been chosen.

Obviously, a parallel argument could be constructed that would show the incoherence of our ideal agent retaining her degree of desire when forced to choose an arbitrary but extremely high speed. (Constructing the argument is left as the proverbial exercise for the reader.) The result seems as unsatisfactory as in the case of certainty. Would the fact that she is confronted with this kind of choice necessarily *alter* her mental states, rendering her either *less* anxious to get to New York or else causing her to make some kind of serious mistake in her practical reasoning?[47] The first option does not seem psychologically inevitable; and the second option seems to be an overly harsh judgment. So the expression principle B**, which seems safe enough in the finite case, should not be expected to remain true when an agent believes she is confronted with an Ever Better Situation.

9 Conclusion

The no-best-world argument may well have teeth even for those who reject consequentialism, despite its utilitarian flavor. It may even survive complications arising from the difficulty of comparing worlds with infinitely many locations of value or locations of infinite value; and the possibility of incomparable worlds. However the engine behind the argument is Rowe's principle B, an expression principle. And reflection upon the application of expression principles to beings facing an infinite (or indefinite) range of options strongly suggests their invalidity in such cases. The possibility of a maximum of some psychological characteristic – certainty, desire, moral goodness, love; or, for that matter, uncertainty, distaste, moral turpitude, hatred – should not be undermined by the possibility of circumstances arising in which the degree of the characteristic cannot be fully

47 As in the parallel argument about certainty, some will say that the traveler in this Ever Better Situation is being perfectly rational, so long as she chooses a speed that is *fast enough*. Others will say that, although it would be positively irrational to choose not to go at all (so it is rationally required to choose some speed or other), nevertheless, whatever speed is chosen, the ideal traveler is open to some kind of rational criticism. (See note 12 above.)

expressed. If some such virtues, vices, and attitudes admit of highest possible degrees (or at least a maximum that is relative to a kind of creature), it should be possible for an agent to have the highest degree and yet make a choice that does not perfectly reflect the character or attitude of the agent – at least, this should be possible when the agent is facing an infinite (or indefinite) range of options. And of course that is exactly the sort of circumstance God is supposed to be in, on the assumptions of the no-best-world argument.[48]

Bibliography

Adams, R. (1987), *The Virtue of Faith*, New York: Oxford University Press.
Beilby, J. K. and Eddy, P. R. (eds.) (2001), Divine Foreknowledge: Four Views, Downers Grove, Illinois: Inter Varsity Press.
Bennett, J. (1988), *Events and their Names*, Indianapolis: Hackett.
Bostrom, N. (2011), "Infinite Ethics", in *Analysis and Metaphysics*: 10, 9–59.
Byron, M. (ed.) (2004), *Satisficing and Maximizing*, Cambridge: Cambridge University Press.
Chisholm, R. M. (1976), *Person and Object*, LaSalle, Illinois: Open Court.
Climenhaga, N. (forthcoming), "Infinite Value and the Best of All Possible Worlds", in *Philosophy and Phenomenological Research*.
Dreier, J. (2004), "Why Ethical Satisficing Makes Sense and Rational Satisficing Doesn't", in Byron (2004), 131–154.
Evans, G. R. (1982), *Augustine on Evil*, Cambridge: Cambridge University Press.
Flint, T. (1998), *Divine Providence*: *The Molinist Account*, Ithaca, NY: Cornell University Press.
Grover, S. (1988), "Why Only the Best Is Good Enough", in *Analysis*: 48, 224.
Hamkins, J. D. and Montero, B. (2000), "With Infinite Utility, More Needn't be Better", in *Australasian Journal of Philosophy*: 78, 231–240.
Hasker, W. (2004), *Providence, Evil and the Openness of God*, London and New York: Routledge.
Howard-Snyder, D. and Howard-Snyder, F. (1994), "How an Unsurpassable Being Can Create a Surpassable World", in *Faith and Philosophy*: 11, 260–268.
Howard-Snyder, D. and Howard-Snyder, F. (1996), "The Real Problem of No Best World", in *Faith and Philosophy*: 13, 422–425.
Korsgaard, Ch. (1983), "Two Distinctions in Goodness", in *Philosophical Review*: 92, 169–195.

48 I have told these parallel stories about certainty and desire in many contexts, and learned a lot from the ensuing discussions. I thank participants in workshops at the École Normale Supérieure; the Baylor Summer Seminar at Columbia, Missouri; the Boston University Institute for Philosophy and Religion; Oxford University's Samuel Butler Society; the God and Goodness workshop hosted by the Rutgers Center for the Philosophy of Religion; and "Quo vadis, Metaphysics?" (a conference honoring Peter van Inwagen) in Warsaw, Poland. I am particularly grateful to Daniel Rubio and Jamie Dreier for some last minute help with my criticism of (B*).

Kraay, K. (2006), "God and the Hypothesis of No Prime Worlds", in *International Journal for Philosophy of Religion*: 59, 49–68.
Kraay, K. (2010), "Theism, Possible Worlds, and the Multiverse", in *Philosophical Studies*: 147, 355–368.
Kretzmann, N. (1990), "A Particular Problem of Creation", in *Being and Goodness*, edited by S. MacDonald, Ithaca, NY: Cornell University Press, 208–249.
Langtry, B. (2008), *God, the Best, and Evil*, Oxford: Oxford University Press.
Lauwers, L. and Vallentyne, P. (2004), "Infinite Utilitarianism: More is Always Better", in *Economics and Philosophy*: 20, 307–330.
Lemos, N. (2015), "A Defense of Organic Unities", in *Journal of Ethics*: 19, 125–141.
Leftow, B. (2005a), "No Best World: Moral Luck", in *Religious Studies*: 41, 165–181.
Leftow, B. (2005b), "No Best World: Creaturely Freedom", in *Religious Studies*: 41, 269–285.
Leftow, B. (forthcoming), "God's Freedom and the Best", in *TheoLogica*.
Monton, B. (2010), "Against Multiverse Theodicies", in *Philo*: 13, 113–135.
Morris, T. V. (1993), "Perfection and Creation", in Stump (1993), 234–247.
Pollock, J. (1983), "How Do You Maximize Expectation Value?", in *Nous*: 17, 409–421.
Pruss, A. R. (2017), "Divine Creative Freedom", in *Oxford Studies in the Philosophy of Religion*, Vol. 7, edited by J. Kvanvig, Oxford: Oxford University Press, 213–238.
Rowe, W. L. (1993), "The Problem of Divine Perfection and Freedom", in Stump (1993), 223–233.
Rowe, W. L. (1994), "The Problem of No Best World", in *Faith and Philosophy*: 11, 269–271.
Rowe, W. L. (2002), "Can God Be Free?", in *Faith and Philosophy*: 19, 405–424.
Rowe, W. L. (2004), *Can God Be Free*?, Oxford: Clarendon Press.
Schmidtz, D. (2004), "Satisficing as a Humanly Rational Strategy", in Byron (2004), 30–58.
Slote, M. (1984), "Satisficing Consequentialism", in *Proceedings of the Aristotelian Society*: 58, 139–163.
Slote, M. (1989), *Beyond Optimizing*, Cambridge, MA: Harvard University Press.
Stump, E. (ed.) (1993), *Reasoned Faith*, Ithaca and London: Cornell University Press.
Swenson, P. (2018), "Normative Principles for No-Best-World", Unpublished manuscript.
Swenson, P. and Zimmerman, D. (eds.) (2019), *Oxford Studies in the Philosophy of Religion*: 9, Oxford: Oxford University Press.
Stump, E. (ed.) (1993), *Reasoned Faith*, Ithaca and London: Cornell University Press.
Tucker, Ch. (2016), "Satisficing and Motivated Submaximization (in the Philosophy of Religion)", in *Philosophy and Phenomenological Research*: 93, 127–143.
Vallentyne, P. and Kagan, S. (1997), "Infinite Value and Finitely Additive Value Theory", in *The Journal of Philosophy*: 94, 5–26.
Van Inwagen, P. (2006), *The Problem of Evil*, Oxford: Oxford University Press.
Zimmerman, D. (2012), "The Providential Usefulness of 'Simple Foreknowledge'", in *Reason, Metaphysics, and Mind*, edited by K. J. Clark and M. Rea, Oxford: Oxford University Press, 174–196.
Zimmerman, D. (2018), "Ever Better Situations and the Failure of Expression Principles", in *Faith and Philosophy*: 35.
Zimmerman, M. J. (2001), *The Nature of Intrinsic Value*, Lanham, MD: Rowman and Littlefield.

Mirosław Szatkowski

Deficiencies of Gödel's Ontological Proof

Abstract: In this paper we prove that: (i) Gödel's axioms do not result in a thesis for the existence of God; and (ii) the thesis for the existence of God is inconsistent with Gödel's axioms.

1 Introduction

The phrase 'Gödel's ontological proof' is a name for (the Austrian, and later American, logician, mathematician, and philosopher) Kurt Gödel's (1906–1978) system of definitions and axioms to prove the necessary existence of an object that possesses all positive properties, but no negative property. This system of definitions and axioms was never published when Gödel was alive, but a brilliant American logician and philosopher, Dana Scott, made it in 1970 the subject of his seminar in Princeton, thanks to which this proof reached a certain small group of logicians. For the first time, H. Sobel in 1987 published a copy of the two-pages text entitled 'Gödel's ontological proof', which had been found in one of Gödel's notebooks.[1] Another version entitled 'Scott's ontological proof', slightly different from Gödel's one, can be also found in Sobel's paper.[2] Both versions contain, in addition to the system of definitions and axioms, sketchy derivations of theorems, in particular the thesis of the necessary existence of God.

Without doubt, Gödel's ontological proof is amongst the most discussed formal proofs for the existence of God in philosophical literature of the last fifty years. It should be, however, added that the question "Whether or not God exists?" was always one of the more mysterious and speculative issues in philosophy. God, as the only omniscient, omnipotent, omnibenevolent – and more generally, possessing all positive properties in the highest degree – being, the efficient cause of all things, a causally inaccessible and unobservable being, etc., can not be discovered using only the tools of empirical sciences. The question is then whether the existence or non-existence of God can be found in the process of more or less complicated reasoning, fulfilling strictly defined conditions

0 The research was funded by the National Science Center, within the research grant OPUS 4, DEC-2012/07/B/HS1/01978.

1 See Sobel (1987). Cf. also Gödel (1995) and Gödel (2004).

2 See Sobel (1987). Cf. also Scott (2004).

https://doi.org/10.1515/9783110664812-025

of deductive correctness. This deductive correctness guarantees the truthfulness of a given belief or sentence by virtue of other beliefs or sentences previously recognized as true. That is, why the most legitimate questions are: Can the existence of God be proven according to the mathematical-logical standard? Can one find in philosophical literature at least one proof for the existence of God which would fulfill all mathematical-logical requirements? For many researchers, especially those fascinated by the mathematical authority of Gödel, it is inconceivable to undermine the formal value of his ontological proof. The reason is that Kurt Gödel is regarded as one of the most prominent mathematician and logicians of all time. For example, the brilliant Hungarian-American mathematician, physicist, inventor, computer scientist, and polymath, John von Neumann called him the greatest logician since Aristotle. Recall here, the greatest achievements of Gödel are: (i). The first incompleteness theorem, which says that in any consistent formal theory T, within which a certain amount of elementary arithmetic can be carried out, there are statements of the language of T which can neither be proved nor disproved in T; (ii). The second incompleteness theorem, an extension of the first, which shows that any consistent theory T, within which a certain amount of elementary arithmetic can be carried out, cannot demonstrate its own consistency; (iii). The proof that the Axiom of Choice and the Continuum Hypothesis are consistent with Zermelo-Frankel set theory; (iv). Contributions to proof theory by clarifying the connections between classical, intuitionistic and modal logics; and (v). Solutions to the equations of general relativity in which there are closed time-like curves – Gödel suggested that this is good reason to suppose that time is not real.

Unfortunately, there is some confusion in philosophical literature. Some authors – for example, Anderson (1990), Anderson and Gettings (1996), Fitting (2002), Fuhrmann (2005), Hájek (1996, 2002a,b, 2011) and Kovač (2003) – use the phrase 'Gödel's ontological proof' for the system of three definitions and five axioms, which Sobel named 'Dana Scott's ontological proof'. It is hard to find a satisfying answer to the question: What is the cause of this confusion? We will keep Sobel's distinction between Godel's and Scott's ontological proof, and we will prove that the first one does not meet the basic deductive requirements, since: (1). Gödel's axioms do not result in a thesis for the existence of God; and (2) The thesis for the existence of God is inconsistent with Gödel's axioms.

2 The presentation of Gödel's ontological proof

2.1 The informal presentation of Gödel's theory.

Gödel uses a kind of modal language with a 2^{nd} order notion of a *positive property* as a primitive, which he introduces with no elaborate clarification. However, his terse and sometimes cryptic explanations yield that he offers two readings of this notion: (1). positive in a *moral - aesthetic* sense. The positiveness in this sense is independent of the *accidental structure of the world* ; and (2). positive in a sense of *pure attribution*. The positiveness in this sense is said to be *opposed to privation*.

The additional three concepts are introduced by the definitions:

Definition 1: A *God* is any being that has every positive property;

Definition 2: A property A is an *essence* of an object x if and only if A entails every property of x;

Definition 3: An object x has the property of *necessarily existing* if and only if its essence is necessarily exemplified.

These above concepts are characterized by the following axioms:

Axiom 1: The conjunction of positive properties is also positive;

Axiom 2: A property or its complement is positive;

Axiom 3: If a property is positive, then its complement is not positive;

Axiom 4: If a property is positive, then it is necessarily positive;

Axiom 5: The property of necessary existence is a positive property;

Axiom 6: Any property entailed by a positive property is positive.

The above set of definitions and axioms was proposed by Gödel with a view to proving, by means of an appropriate modal logic of the 2^{nd} order, that:

Theorem: A God necessarily exists.

2.2 The formal presentation of Gödel's theory.

A language capable of expressing Gödel's axioms should be equipped with a 2^{nd} order unary predicate P, where $P(\alpha)$ is to be read: *the property α is positive*, a necessity symbol L, two sorts of variables: $x, y, z, \ldots$ (1^{st} order), $\alpha, \beta, \psi, \ldots$ (2^{nd} order),

Boolean operators: $\cap$, – (intersection and complementation), customary logical symbols such as: $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, $\neg$ (conjunction, disjunction, implication, biconditional, negation) and quantifiers $\forall$, $\exists$ for both sorts of variables.

Gödel's theory is based on the following set of definitions:

(2.1) $\mathsf{G}(x) \stackrel{df}{=} \forall\alpha(P(\alpha) \rightarrow \alpha(x))$
$\mathsf{G}(x)$ is read: *x is God-like* or simply *x is a God*,

(2.2) $\alpha\ \mathsf{Ess}\ x \stackrel{df}{=} \forall\beta(\beta(x) \rightarrow \mathsf{L}\forall y(\alpha(y) \rightarrow \beta(y)))$
$\alpha\ \mathsf{Ess}\ x$ is read: *a property α is an essence of entity x*,

(2.3) $\mathsf{NE}(x) \stackrel{df}{=} \forall\alpha(\alpha\ \mathsf{Ess}\ x \rightarrow \mathsf{L}\exists y\alpha(y))$
$\mathsf{NE}(x)$ is read: *x necessarily exists*,

and axioms:

(2.4) $P(\alpha) \wedge P(\beta) \rightarrow P(\alpha \cap \beta)$,

(2.5) $P(\alpha) \vee P(-\alpha)$,

(2.6) $P(\alpha) \rightarrow \neg P(-\alpha)$,

(2.7) $P(\alpha) \rightarrow \mathsf{L}P(\alpha)$,

(2.8) $P(\mathsf{NE})$,

(2.9) $P(\alpha) \wedge \mathsf{L}\forall x(\alpha(x) \rightarrow \beta(x)) \rightarrow P(\beta)$.

3 The unprovability of God's existence with Gödel's axioms

The result announced in the title of this section is proved by the following reasoning:

It is clear that the above axioms of K. Gödel leave a degree of freedom in interpreting the necessity symbol L. If, however, our theory meets the following rather natural condition:

– no new formula *not* containing the symbol L can be proved if the axiom $\mathsf{L}\phi \leftrightarrow \phi$ is added to the theory

then the axioms: (2.4) – (2.9) are too weak to prove the much needed sentence: $\exists x \mathsf{G}(x)$. To see this, assume that variables of 1^{st} order range over the set of natural numbers ω, variables of 2^{nd} order range over 2^{ω} and P is interpreted as a non-principal ultrafilter of the Boolean algebra 2^{ω} containing all co-finite sets of natural numbers. Then, assuming that $\mathsf{L}\phi \leftrightarrow \phi$, for every formula ϕ, we get that the axioms (2.4) – (2.9) are satisfied and the sentence $\exists x \mathsf{G}(x)$ is false because the intersection of all co-finite sets is empty.

Describing this reasoning, we indicated a model in which Gödel's axioms are true, and the thesis of the existence of God is false. In this way, we demonstrated that the set of axioms of Gödel's theory is consistent – in contrary to the statement of Christoph Benzmüller and Bruno Woltzenlogel Paleo about the inconsistency of Gödel's axioms, see Benzmüller and Paleo (2016).

4 The inconsistency of the thesis of God's existence with Gödel's axioms

We need a definition of an empty set, which is as follows:

(4.1) $\varnothing(x) \stackrel{\mathrm{df}}{=} (\alpha(x) \wedge \neg\alpha(x))$,

and the following formulas:

(4.2) All what is needed for classical propositional logic,

(4.3) $\forall\xi(\phi \rightarrow \psi) \rightarrow (\forall\xi\phi \rightarrow \forall\xi\psi)$,

(4.4) $\forall\alpha\phi \rightarrow \phi(\alpha/A)$ where A is a term of the 2^{nd} sort,

(4.5) $\forall x(-\alpha(x) \leftrightarrow \neg\alpha(x))$,

(4.6) $\mathsf{L}(\phi \rightarrow \psi) \rightarrow (\mathsf{L}\phi \rightarrow \mathsf{L}\psi)$,

(4.7) $\mathsf{L}\phi \rightarrow \phi$.

Axioms of the theory **GO** are defined as follows:

(4.8) (i) Each formula (2.4)–(2.9) and (4.2)–(4.7) is an axiom of the theory **GO**;
(ii) $Q\phi$ is an axiom of the theory **GO** of ϕ is its axiom, where Q is a finite (including the empty) sequence of the symbols each of which has the form $\forall\xi$ or L.

THE PROOF OF THE FORMULA $\bot(\stackrel{\mathrm{Df}}{=} \mathit{FALSUM})$ FROM THE AXIOMS OF THE THEORY **GO** AND THE GÖDEL'S THEOREM (**GTH**): $\mathsf{L}\exists x\mathsf{G}(x)$:

1.	$\mathsf{L}\forall x\forall\alpha\mathsf{L}\forall y(\varnothing(y) \to \alpha(y))$	(4.2), (4.8)
2.	$\mathsf{L}\forall x\forall\alpha(\alpha(x) \to \mathsf{L}\forall y(\varnothing(y) \to \alpha(y)))$	(4.2), (4.8), (4.1), (4.5), **1**, MP
3.	$\mathsf{L}\forall x(\varnothing\ \mathsf{Ess}\ x)$	(2.2), (4.8), (4.3), (4.6), **2**, MP
4.	$\mathsf{L}\exists x(\varnothing\ \mathsf{Ess}\ x) \to \mathsf{L}\forall x(\varnothing\ \mathsf{Ess}\ x)$	(4.2), **3**, MP
5.	$\mathsf{L}\exists x(\varnothing\ \mathsf{Ess}\ x) \to \mathsf{L}\forall x(\varnothing\ \mathsf{Ess}\ x \wedge \mathsf{M}\forall y\neg\varnothing(y))$	(4.2), (4.8), (4.3), (4.6), **4**, MP
6.	$\mathsf{L}\exists x(\varnothing\ \mathsf{Ess}\ x) \to \mathsf{L}(\mathsf{M}\forall y\neg\varnothing(y) \to \forall x(\varnothing\ \mathsf{Ess}\ x \wedge \mathsf{M}\forall y\neg\varnothing(y)))$	(4.2), (4.8), (4.3), (4.6), **5**, MP
7.	$\mathsf{L}\exists x(\varnothing\ \mathsf{Ess}\ x) \to (\mathsf{L}\exists x(\varnothing\ \mathsf{Ess}\ x \to \mathsf{L}\exists y\varnothing(y)) \to \mathsf{LL}\exists y\varnothing(y))$	(4.2), (4.8), (4.3), (4.6), **6**, MP
8.	$\mathsf{L}\exists x(\varnothing\ \mathsf{Ess}\ x \to \mathsf{L}\exists y\varnothing(y)) \to \mathsf{LL}\exists y\varnothing(y)$	**3**, **7**, MP
9.	$\mathsf{L}\forall x(\mathsf{G}(x) \to \forall\alpha(P(\alpha) \to \alpha(x)))$	(2.1), (4.8)
10.	$\mathsf{L}\forall x(\mathsf{G}(x) \to (P(\mathsf{NE}) \to \mathsf{NE}(x)))$	(4.2), (4.8), (4.4), (4.3), (4.6), **9**, MP
11.	$\mathsf{L}\forall x(\mathsf{G}(x) \to \mathsf{NE}(x))$	(4.2), (4.8), (4.3), (4.6), (2.8), (2.7), (4.4), **10**, MP
12.	$\mathsf{L}\exists x\mathsf{G}(x) \to \mathsf{L}\exists x\mathsf{NE}(x))$	(4.2), (4.8), (4.3), (4.6), **11**, MP
13.	$\mathsf{L}\exists x\mathsf{NE}(x)$	**GTh**, **12**, MP
14.	$\mathsf{L}\exists x\forall\alpha(\alpha\ \mathsf{Ess}\ x \to \mathsf{L}\exists y\alpha(y))$	(2.3), (4.8), (4.3), (4.6), **13**, MP
15.	$\mathsf{L}\exists x(\varnothing\ \mathsf{Ess}\ x \to \mathsf{L}\exists y\varnothing(y))$	(4.4), (4.8), (4.3), (4.6), **14**, MP
16.	$\mathsf{LL}\exists y\varnothing(y)$	**8**, **15**, MP
17.	$\exists y\varnothing(y)$	**16**, (4.7), MP
18.	$\bot$	**17**

5 Wroński's suggestion

Finally, I would like to share a suggestion concerning proofs for the existence of God, received from Professor Andrzej Wroński (a Polish logician). Here it is:

> Proving the existence of God seems to be similar to proving the non-contradiction of arithmetic. We believe that arithmetic is non-contradictory. Yet, as it constitutes the foundation of our entire knowledge, we would also like to have a convincing proof in addition to the belief. It is known (2^{nd} Gödel's Theorem) that the proof for the non-contradiction of arithmetic

> cannot be successful without referring to external means (not formalized in arithmetic). Hence, the non-contradiction of the theory which "is able" to prove the non-contradiction of arithmetic is as problematic as the non-contradiction of arithmetic itself. But for any predetermined natural number n, the fact that there is no contradiction among arithmetic statements whose formal proofs' length is $\leq n$ can be easily proved by means formalized in arithmetic (reflection principle). Is there any theological counterpart of the reflection principle? It may be useful to differentiate between the global and local proof for the existence of God, show the possibility of local proof, impossibility of global proof, etc.

We hope that this Wroński's suggestion will find a response in philosophical literature and become a subject of research.

Bibliography

Anderson, C. A.(1990), "Some Emendations of Gödel's Ontological Proof", in *Faith and Philosophy*: 7, 291–303.

Anderson, C. A. and Gettings, M.(1996), " Gödel's Ontological Proof Revised", in *Gödel'96: Logical Foundations of Mathematics, Computer Science, and Physics*. *Lecture Notes in Logic*, Springer: 6, 167–172.

Benzmüller, C. and Paleo, B. W.(2016), "The Inconsistency in Gödel's Ontological Argument: A Success Story for AI in Metaphysics", in *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence – IJCAI*, edited by S. Kambhampati, AAAI Press.

Fitting, M. (2002), *Types, Tableaus, and Gödel's God*, Kluwer Academic Publishers.

Fuhrmann, A. (2005), "Existenz und Notwendigkeit – Kurt Gödels axiomatische Theologie", in *Logik in der Philosophie*, edited by W. Spohn, Heidelberg, 349–374.

Gödel, K. (1995), "Ontological Proof", in *Kurt Gödel Collected Works*, edited by S. Feferman, J. W. Dawson Jr., W. Goldfarb, C. Parsons, and R. M. Soloway, vol. III, New York and Oxford: Oxford University Press,403–404.

Gödel, K. (2004), "Notes in Kurt Gödel's Hand", in *Logic and Theism*: *Arguments For and Against Beliefs in God*, authored J. H. Sobel, New York: Cambridge University Press,144–145.

Hájek, P. (1996), "Magari and others on Gödel's Ontological Proof", in *Logic and Algebra*, edited by A. Ursini and P. Agliano, Marcel Dekker, Inc., 125–135.

Hájek, P. (2002a), "Emendation of Gödel's Ontological Proof", in *Studia Logica*: 71(2), 149–164.

Hájek, P. (2002b), "Der Mathematiker und die Frage der Existenz Gottes (betreffend Gödels ontologischen Beweis)", in *Kurt Gödel. Wahrheit und Beweisbarkeit*, Part II. Kompedium zum Werk, edited by B. Buldt, E. Köhler, M. Stöltzner, P. Weibel, C. Klein, W. Depauli-Schimanowich-Göttig, Wien: ÖBV et HPT, 325–335.

Hájek, P. (2011), "Gödel's Ontological Proof and its Variants", in *Kurt Gödel and the Foundations of Mathematics*: *Horizons of Truth*, edited by M. Baaz, Cambridge: Cambridge University Press, 307–322.

Kovač, S. (2003), "Some Weakened Gödelian Ontological Systems", in *Journal of Philosophical Logic*: 32, 565–588.

Scott, D. (2004), "Notes in Dana Scott's Hand", in *Logic and Theism: Arguments For and Against Beliefs in God*, authored by J. H. Sobel, New York: Cambridge University Press, 145–146.

Sobel, J. H. (1987), "Gödel's Ontological Argument", in *On Being and Saying: Essays for Richard Cartwright*, edited by J. J. Thomson, Cambridge MA: MIT Press, 241–261.

Authors of Contributed Papers

Andrea C. Bottani is Professor of Theoretical Philosophy at the University of Bergamo (Italy) and invited Professor at the University 'San Raffaele' of Milan and at the FTL of Lugano (CH). He taught and did research at the Universities of Genova, Fribourg, Neuchâtel and Urbino. He was Senior Research Fellow at King's College London, Visiting Scholar at the Columbia University New York, President of the Italian Society for Analytic Philosophy, Member of the Steering Committee of the Italian Society for Logic and Philosophy of Sciences and a founder of the Italian Society for Theoretical Philosophy. He is currently member of the Board of Teachers of the Doctorate School "F.I.N.O." (Northwestern Italian Philosophy Consortium) in Philosophy. Among his publications are two authored books, nine edited volumes and more than sixty book chapters and journal articles. His area of specialization includes metaphysics, ontology and philosophy of language.

Christopher Daly is a Professor of Philosophy at the University of Manchester. His principal research interests are in metaphysics, philosophical methodology, and philosophy of language. He is the author of *Introduction to Philosophical Methods* (Broadview, 2010) and *Philosophy of Language*: *An Introduction* (Bloomsbury, 2012). He has published papers on natural kinds, properties, truthmaker theory, grounding, philosophy of mathematics, and fictionalism. He is an Associate Editor of the *Australasian Journal of Philosophy*. He and David Liggins are currently directing a research project entitled *The Foundations of Ontology*, funded by the UK Arts and Humanities Research Council.

William Jaworski is Associate Professor of Philosophy at Fordham University in New York City. He is the author of Structure and the Metaphysics of Mind: How Hylomorphism Solves the Mind-Body Problem (Oxford University Press 2016) and Philosophy of Mind: A Comprehensive *Introduction* (Wiley-Blackwell 2011). He works on topics in metaphysics, philosophy of mind, and the philosophy of religion.

Christian Kanzian is a Professor of Philosophy at the University of Innsbruck, Austria. He is, since 2006, the president of Austrian Ludwig Wittgenstein Society (Österreichische Ludwig Wittgenstein Gesellschaft). Kanzian's main interests are the analytic philosophy, the history of philosophy, and ontology. He is the author of four books: *Originalität und Krise – Zur* systematischen Rekonstruktion der Frühschriften Kants (1994), Grundproblerne der Analytischen Ontologie (with E. Runggaldier, 1998), Ereignisse und andere Partikularien (2001), Ding – Substanz – Person. Eine Alltagsontologie (2009), and an editor or co-editor of eleven books. He has also published over seventy articles.

Øystein Linnebo received an M.A. in mathematics in 1995 at the University of Oslo and a Ph.D. in Philosophy in 2002 from Harvard. He is currently Professor of Philosophy at the University of Oslo, where he works in philosophical logic, philosophy of mathematics, metaphysics, and early analytic philosophy (especially Frege). He has published more than 50 scientific articles and is the author of two books, Philosophy of Mathematics (Princeton University Press, 2017) and Thin Objects: An Abstractionist Account (Oxford University Press, 2018).

Anna-Sofia Maurin is Professor of Theoretical Philosophy at the University of Gothenburg, working in metaphysics and metametaphysics. Her main areas of expertise are theories of properties (especially tropetheory), infinite regress arguments (especially Bradley's regress) and metaphysical explanation. She is the author of *If Tropes* (Kluwer 2002) and is presently the principal investigator of a major research project on metaphysical explanation financed by Riksbankens Jubileumsfond.

Uwe Meixner is a professor of philosophy at the University of Augsburg, Germany. His main areas of philosophical interest are metaphysics, logic, and philosophical theology. He has published four monographs in English: The Two Sides of Being (2004), The Theory of Ontic Modalities (2006), Modelling Metaphysics (2010), and Defending Husserl (2014).

Kevin Mulligan is ordinary professor of philosophy at the University of Italian Switzerland, Lugano; Director of Research, Institute of Philosophical Studies, Faculty of Theology, Lugano; honorary professor of analytic philosophy, Geneva. He has written about analytic metaphysics, the philosophy of mind and the history of Austro-German philosophy from Bolzano to Wittgenstein. He is a member of the Swedish and European Academies.

Joanna Odrowąż-Sypniewska is a Professor of Philosophy at the University of Warsaw. She received a Ph.D. in Philosophy form University of Warsaw (1999) and from St Andrews University (2001). Her main area of interest is philosophy of language (vagueness, natural kind terms, semantic minimalism and contextualism debate). She has published three books in Polish and many papers (in English and in Polish).

Francesco Orilia studied philosophy at the University of Palermo and at Indiana University, where he obtained his Ph.D., with Hector-Neri Castañeda as supervisor, in 1986. He is a professor in the Philosophy and Human Sciences Section of the University of Macerata's Department of Humanities. His main research interests are in these areas: logic (logical paradoxes, normativity of logic), philosophy of language (reference), philosophy of mind (mindbody problem, mental causation), philosophy of time, ontology (properties and relations, relational order, Bradley's regress). In addition to many papers on these topics, he has published in English the monographs Predication, Analysis and Reference (CLUEB, 1999), and Singular Reference. A Descriptivist Perspective (Springer, 2010), and coedited several collections, including, with M. Paolini Paoletti, Philosophical and Scientific Perspectives on Downward Causation (Routledge, 2017).

Carl J. Posy is Professor Emeritus of Philosophy at the Hebrew University of Jerusalem. His published work covers philosophical logic, philosophy of logic and of mathematics, and the history of philosophy, concentrating on Kant and his predecessors.

Graham Priest is Distinguished Professor of Philosophy at the Graduate Center, City University of New York, and Boyce Gibson Professor Emeritus at the University of Melbourne. He is known for his work on non-classical logic, particularly in connection with dialetheism, on the history of philosophy, and on Buddhist philosophy. He has published articles in nearly every major philosophy and logic journal. His books include: In *Contradiction*: *A Study of the Transconsistent*, Martinus Nijhoff 1987 (2nd edition: Oxford University Press 2006); *Beyond the Limits of*

Thought, Cambridge University Press 1995 (2nd edition: Oxford University Press 2002); Towards Non-Being: the Semantics and Metaphysics of Intentionality, Oxford University Press 2005 (2nd edition, 2016); *Doubt Truth to be a Liar*, Oxford University Press 2006; Introduction to Non-Classical Logic From If to Is, Cambridge University Press 2008; One, Oxford University Press 2014; *The Fifth Corner of Four*, Oxford University Press, forthcoming. When not doing philosophy, he likes to do philosophy. Further details can be found at: grahampriest.net.

Nicholas Rescher is Distinguished University Professor of Philosophy at the University of Pittsburgh. In a productive research career extending over six decades he has well over one hundred books to his credit. Fourteen books about Rescher's philosophy have been published in five languages. He has served as a President of the American Philosophical Association, of the American Catholic Philosophical Association, of the American G. W. Leibniz Society, of the C. S. Peirce Society, and of the American Metaphysical Society as well as Secretary General of the International Union of History and Philosophy of Sciences. Rescher has been elected to membership in the American Academy of Arts and Sciences, the Academia Europea, the Royal Society of Canada, and the Royal Asiatic Society of Great Britain. He has been awarded the Alexander von Humboldt prize for Humanistic Scholarship in 1984, the Belgian Prix Mercier in 2005, the Aquinas Medal of the American Catholic Philosophical Association in 2007, the Founder's Medal of the Metaphysical Society of America in 2016, and the Helmholtz Medal of the Germany Academy of Sciences (Berlin/Brandenburg) in 2016. In 2011 he was awarded the premier cross of the Order of Merit (Bundesdienstkreuz Erster Klasse) of the Federal Republic of Germany, and honorary degrees have been awarded to him by eight universities on three continents.

Benjamin Schnieder holds a Chair for Theoretical Philosophy at the University of Hamburg, where he is also the director of the research group Phlox. His areas of specialization are metaphysics and the philosophy of language and logic; he has a particular interest in the semantics of ground-theoretical vocabulary and the metaphysics of grounding.

Stewart Shapiro received an M.A. in mathematics in 1975, and a Ph.D. in philosophy in 1978, both from the State University of New York at Buffalo. He is currently the O'Donnell Professor of Philosophy at The Ohio State University, and he serves as Distinguished Visiting Professor at the University of Connecticut, and as Professorial Fellow at the University of Oslo. He has contributed to the philosophy of mathematics, philosophy of language, logic, and philosophy of logic, publishing monographs on higher-order logic, structuralism, vagueness, and pluralism in logic.

Gila Sher is a Professor of Philosophy at the University of California, San Diego. She is working on foundational issues in epistemology, truth, and logic. Her books include: *The Bounds of Logic*: A Generalized Viewpoint (MIT, 1991) and Epistemic Friction: An Essay on Knowledge, Truth, and *Logic* (Oxford, 2016). A few of her papers are: "Ways of Branching Quantifiers" (*Linguistics and Philosophy* 1990), "Did Tarski Commit Tarski's Fallacy?" (*Journal of Symbolic Logic* 1996), "Is There a Place for Philosophy in Quine's Theory?" (*Journal of Philosophy* 1999), "The Formal-Structural View of Logical Consequence" (*Philosophical Review* 2001), "Epistemic Friction: Reflections on Knowledge, Truth, and Logic" (*Erkenntnis* 2010), "Is Logic in the Mind or in the World?" (*Synthese* 2011), "The Foundational Problem of Logic" (*Bulletin of Symbolic Logic*

2013), "Truth and Scientific Change" (*Journal of General Philosophy of Science* 2017), "Lessons on Truth from Kant" (*Analytic Philosophy* 2017). She is a co-editor of *Between Logic and* Intuition: Essays in Honor of Charles Parsons (Cambridge 2000).

Peter Simons is Emeritus Professor of Philosophy at Trinity College Dublin. He specialises in metaphysics and ontology, pure and applied; the philosophy of logic and mathematics; and the history of philosophy in Central Europe in the 19th and 20th centuries. He is the author or co-author of 5 books and some 300 articles. He is a member of the British, Irish, European and Polish Academies.

Eleonore Stump is the Robert J. Henle Professor of Philosophy at Saint Louis University, where she has taught since 1992. She is also Honorary Professor at Wuhan University and at the Logos Institute, St.Andrews, and she is a Professorial Fellow at Australian Catholic University. She has published extensively in philosophy of religion, contemporary metaphysics, and medieval philosophy. Her books include her major study Aquinas (Routledge, 2003) and her extensive treatment of the problem of evil, Wandering in Darkness: Narrative and the Problem of Suffering (Oxford, 2010) and her Atonement (Oxford, 2018). She has given the Gifford Lectures (Aberdeen, 2003), the Wilde lectures (Oxford, 2006), the Stewart lectures (Princeton, 2009), and the Stanton lectures (Cambridge, 2018). She is past president of the Society of Christian Philosophers, the American Catholic Philosophical Association, and the American Philosophical Association, Central Division; and she is a member of the American Academy of Arts and Sciences.

Mirosław Szatkowski is a professor of philosophy at the Warsaw University of Technology, Poland, and the Director of the International Center for Formal Ontology. He earned his PhD in philosophy from Jagiellonian University in Cracow, Poland, and was habilitated at the Ludwig-Maximilians University in Munich, Germany. Szatkowski's main fields of research are: logic, the foundations of mathematics, and formal ontology. In these areas, he has published papers in the following professional journals: Studia Logica, Zeitschrift für mathematische Logik und Grundlagen der Mathematik (Mathematical Logic Quarterly), Archiv für Mathematische Logik und Grundlagenforschung (Archive for Mathematical Logic), Notre Dame Journal of Formal Logic, Journal of Applied Non-Classical Logics, Journal of Logic, Language and Information, and *Metaphysica*; and in several collective volumes. He has edited five volumes: *Ontological Proofs* Today (Frankfurt: Ontos Verlag, 2012), Dualistic Ontology of the Human Person (München: Philosophia, 2013), *Substantiality ana Causality* (Boston/Berlin/Munich: Walter de Gruyter, 2014), *God, Truth, and other Enigmas* (Boston/Berlin/Munich: Walter de Gruyter, 2015), and *Analytically Oriented Thomism* (Editiones Scholasticae, 2016).

Alfredo Tomasetta is a tenure-track researcher at the University School for Advanced Studies IUSS of Pavia, Italy. He specializes in philosophy of mind and metaphysics. In addition to these primary areas of research, he is also interested in philosophical logic, philosophy of language, and classical Indian philosophy. He has published four books in Italian and many papers (in Italian and in English).

Peter van Inwagen is an American analytic philosopher and the John Cardinal O'Hara Professor of Philosophy at the University of Notre Dame. He previously taught at Syracuse University and

earned his PhD from the University of Rochester under the direction of Richard Taylor and Keith Lehrer. Van Inwagen is one of the leading figures in contemporary metaphysics, philosophy of religion, and philosophy of action. He has published nine books and over 200 articles on general metaphysics, free will, material objects and human persons, philosophy of religion, logic and language. In 2005 Peter van Inwagen was elected to the American Academy of Arts and Sciences, and in 2011 he received the degree Doctor Divinitatis (honoris causa) from the University of St Andrews.

Takashi Yagisawa is Professor of Philosophy at California State University, Northridge. After receiving his BA in philosophy from University College, London, he earned his PhD in philosophy from Princeton University under the supervision of David Lewis. His principal research interests are in metaphysics, especially modal metaphysics and its applications. He has published articles in metaphysics, philosophy of language, philosophical logic, and philosophy of mind. His books include one in English, *Worlds and Individuals, Possible and Otherwise* (Oxford University Press, 2010), and seven in Japanese, *Introduction to Analytic* Philosophy (Tokyo: Kodansha, 2011), Meaning, Truth, and Existence: Intermediate Introduction to Analytic Philosophy (Kodansha, 2013), From God to Possible Worlds: Advanced Introduction to Analytic Philosophy (Kodansha, 2014), Analytic Philosophy of Alice in *Wonderland* (Kodansha, 2016), *Analyzing* "*Correct*" (Tokyo, Iwanami Shoten, 2016), *Analyzing Logic* (Iwanami Shoten, 2018), and *Analyzing Natural Numbers* (Iwanami Shoten, 2018).

Dean Zimmerman earned a bachelors degree from Minnesota State University–Mankato, and a Ph.D. in philosophy from Brown University. He has taught at the University of Notre Dame, Syracuse University, and Rutgers University, where he is now a professor in the philosophy department and co-director (with Brian Leftow) of the Rutgers Center for the Philosophy of Religion. Zimmerman is founding editor of *Oxford Studies in Metaphysics* (now co-edited with Karen Bennett), and co-editor of *Oxford Studies in the Philosophy of Religion*. He has co-edited several other books, including *Metaphysics*: *The Big Questions* (Blackwell, 2008), *The Oxford Handbook of Metaphysics* (Oxford University Press, 2003), *Contemporary* Debates in Metaphysics (Blackwell, 2008), Persons: Human and Divine (Oxford University Press, 2007), and *God in an Open Universe* (Pickwick, 2011). His publications include over 50 articles in scholarly journals and books.

Person Index

Adams R. 444, 448, 454, 467
Ameseder R. 186, 188, 200
Anderson C. A. 475
Anderson J. O. 238, 245
Anscombe G. E. 217, 222, 229
Aquinas Thomas 425, 426, 428
Aristotle 63, 101, 351, 363, 386
Armstrong D. M. 93, 104, 106, 117, 297, 312
Audi P. 111, 117
Augustine 416
Austin J. L. 162, 180
Azzouni J. 159, 180

Bach E. 155, 180
Baker A. 402, 411
Barnes E. 52, 56
Bechtel W. 257, 258, 263
Beilby J. K. 451, 467
Benacerraf P. 45, 56
Bennett J. 69, 81, 467
Benzmüller C. 475
Bergmann G. 105, 117, 184, 200
Berto F. 316, 328
Bierce A. 83, 101
Billicsich F. 386, 396
Black M. 222, 229, 310, 312
Blatti S. 265, 275
Bliss R. 40, 56
Bolduc E. M. 33
Bolzano B. 62, 63, 65, 81, 184, 188, 191, 201
Bostrom N. 457, 467
Bottani A. C. 297
Brentano F. 183–185, 191, 197, 201
Bricker P. 408, 411
Burke M. 249, 263
Burley W. 387
Byron M. 467

Campbell N. A. 252, 263
Cantor G. 351, 363
Carlson G. 169, 181
Carnap R. 18, 30, 37, 56, 189, 191, 201
Casati F. 229
Casati R. 166, 181
Castañeda H. -N. 96, 97, 101, 288, 296
Chalmers D. J. 251, 263, 405, 411
Chang R. 47, 56
Chierchia G. 169, 181
Chisholm R. M. 21, 446, 467
Climenhaga N. 456, 458–460, 467
Cocchiarella N. 121, 122, 137, 147
Coffa A. 50, 56
Correia F. 62, 70, 74, 81
Cowling S. 308, 312
Craver C. F. 257, 258, 263
Cresswell M. J. 134, 160, 181
Crusius Ch. 60, 81
Cummins R. 56, 257, 258, 263

Daly Ch. 399
Dasgupta S. 56, 112, 117
Dawkins R. 238, 245
De Grazia D. 265, 275
Denkel A. 52, 56
Dennett D. C. 263
deRosset L. 111, 117
Devitt M. 297, 312
Doyle A. C. 373, 396
Dreier J. 467
Dretske F. 166, 181
Dummett M. 49, 56
Dupré J. 244, 245

Eckhart M. 421, 422, 424, 425, 429, 430, 440
Eddy P. R. 451, 467
Evans G. R. 207, 215, 446, 467
Evnine S. J. 250, 263

Farrell J. 402, 411
Feldman R. 410, 411
Feynman R. 19, 20
Finch A. 81

https://doi.org/10.1515/9783110664812-026

Fine K. 38–41, 43, 56, 65, 70, 74, 81, 111, 117, 154, 155, 164, 165, 169, 171, 181, 250, 251, 263
Fitting M. 475
Flint T. 452, 467
Fodor J. 258, 263
Folse H. J. 238, 245
Frege G. 183, 189, 221, 222, 312
Friedman M. 49, 56
Fuhrmann A. 475
Funkhouser E. 308, 312

Gödel K. 475
Gärdenfors P. 308, 312
Gale R. M. 368, 396
Gauss K. F. 351, 352, 363
Geach P. T. 181, 183, 222, 229, 330, 332, 348
Gettings M. 475
Glennan S. 257, 263
Glymour C. 56
Goodwin B. 238, 245
Grandy R. 129, 147
Grattan-Guinness I. 50, 56
Grene M. 252, 263
Grossmann R. 184, 201
Grover S. 443, 467

Hájek P. 475
Haack S. 56
Hacker P. M. S. 160, 164, 181
Hacking I. 312
Haefliger G. 196, 201
Haldane J. 31
Hamkins J. D. 453, 467
Hankisson R. J. 388, 396
Harman G. 45, 56
Hasker W. 461, 467
Hausmann M. 81
Hawley K. 164, 181
Hawthorne J. 148, 308, 312
Heidegger M. 183, 186, 187, 191, 192, 194, 195, 201, 219–222, 227
Heil J. 257, 263, 323, 328
Heller M. 181
Herron M. D. 238, 245
Hochberg H. 287, 288, 296
Horgan T. 343, 348
Hornsby J. 52, 56
Howard-Snyder D. 461, 467
Howard-Snyder F. 461, 467
Hughes Ch. 30, 310, 312
Husserl E. 88, 95, 101, 184, 186–189, 191–194, 197, 200, 201

Ingarden R. 187, 188, 193, 196, 198, 201

James W. 185, 191, 195, 201
Jantzen D. H. 238, 245
Jaworski W. 247, 250, 255, 258, 263
Johnston M. 250, 264–267, 269, 275, 332, 333, 348

Kölbel M. 346, 348
Künne W. 63, 81
Kagan S. 453, 456, 468
Kant I. 18, 37, 57, 62, 81, 143, 148, 392, 396
Kanzian Ch. 315, 323, 328
Kappes Y. 66, 81
Katz G. 169, 181
Kelly T. 399, 411
Kennedy Ch. 347, 348
Kenny A. 183, 197, 201
Kim J. 181
King N. L. 399, 409, 412
Kleene S. C. 359, 363
Kluback W. 221, 229
Koons R. 250, 264
Korsgaard Ch. 444, 467
Koslicki K. 250, 254, 264
Kovač S. 475
Krämer S. 69, 81
Kraay K. 460, 462, 468
Kretzmann N. 31, 388, 396, 461, 468
Kriegel U. 201
Kripke S. 99, 207, 208, 215
Kuhne R. L. 386, 396

Ladyman J. 251, 264
Langton R. 145, 148
Langtry B. 444, 454, 461, 468
Lauwers L. 453, 468
Leśniewski S. 234

Lear J. 351, 363
Leftow B. 451, 454, 461, 468
Leibniz G. W. 59, 60, 62, 71, 75, 81, 309, 312, 367, 368, 370, 373, 374, 376, 389, 391, 392, 394, 396
Leighton R. B. 20
Lemos N. 452, 468
Leslie J. 386, 396
Lewis D. 14, 15, 79, 81, 86, 93, 101, 105, 106, 117, 120, 134, 135, 148, 164, 181, 310, 312, 329, 330, 332, 348, 401, 403, 404, 410–412
Lilly R. 221, 229
Lim J. 268, 275
Linnebo Ø. 52, 57, 351, 354, 363
Locke J. 310, 312
Lopez de Sa D. 330, 332, 334–337, 348
Loux M. J. 103, 105, 108, 118
Lowe E. J. 320, 322, 323, 327, 328, 348
Lycan W. G. 258, 264, 405, 412

MacBride F. 287, 296
Maienborn C. 169, 174, 181
Margulis L. 238, 245
Marmodoro A. 250, 264
Marty A. 183, 184, 186, 189, 195, 196, 201
Matthews G. 449
Maurin A. -S. 103
Maynard-Smith J. 239, 245
McCord A. M. 413, 440
McDaniel K. 155, 157, 181, 183, 187, 202
McGinn C. 157, 160, 178, 181
McGuiness B. 218, 229
McKinnon N. 334, 349
McTaggart J. M. E. 393, 394
Meinong A. 84, 86, 101, 183, 185–187, 189, 192, 201
Meixner U. 83, 101
Merlo G. 331, 333, 338–340, 349
Merricks T. 181, 265, 275
Michon C. 31
Miller B. 157, 181
Moltmann F. 153, 155, 159, 162, 169, 171, 174, 181
Montero B. 453, 467
Monton B. 460, 468
Moore G. E. 200, 202, 452
Moore G. H. 50, 57
Morris G. 59, 81
Morris T. V. 461, 468
Mulligan K. 183, 184, 197, 202, 323, 328
Murphy M. 31
Muyskens R. 155, 182

Nadelhoffer T. 59, 81
Nagel T. 435, 440
Nahmias E. 59, 81
Nolan D. 105, 118
Noonan H. 349
Novak L. 148
Novotny D. 148

O'Callaghan J. 31
Oderberg D. S. 250, 251, 264
Odrowąż-Sypniewska J. 329, 330, 342, 345, 349
Olson E. T. 264, 265, 275
Orłowska E. 343–345, 349
Orilia F. 279, 285, 287, 296

Paleo B. W. 475
Parsons T. 157, 182
Paul L. 106, 118, 142, 148
Pawlak Z. 349
Pearce K. L. 234, 245
Pears D. 218, 229
Pepper J. W. 238, 245
Philosophus J. 84, 86, 87, 90, 91, 93
Pinsent A. 436, 440
Plantinga A. 18, 251, 264, 413, 440
Plato 386
Plebani M. 316, 328
Poli R. 148
Pollock J. 465, 468
Posy C. J. 119, 137, 139, 143, 148
Priest G. 155, 157, 182, 217, 218, 220, 221, 224–229
Pritchard D. 409, 412
Prokop S. 148
Pruss A. R. 368, 396, 454, 459, 468
Przywara E. 193, 202
Puivet R. 31
Putnam H. 14, 23, 132–134, 137, 139, 140, 145, 148

Queller D. 238, 245
Quine W. V. O. 23, 45, 49, 57, 90, 101, 109, 113, 116, 118, 264, 297, 312, 316, 317, 330, 337, 338, 349
Quinton A. 297, 312

Raven M. J. 66, 81, 111, 118
Rea M. C. 148, 250, 264
Rechter O. 148
Reinach A. 192, 202
Rescher N. 74, 81, 367, 386, 394, 396
Richard S. 186, 196, 202
Richardson R. 257, 263
Rodriguez-Pereyra G. 64, 73, 81, 82
Rosen G. 38–41, 57, 65, 74, 82
Roski S. 69, 81
Ross D. 251, 264
Ross J. F. 69, 82
Roughgarden J. 238, 245
Rowe W. L. 69, 70, 82, 368, 396, 443, 444, 447–449, 451, 463, 468
Rudder-Baker L. 322, 328
Rumfitt I. 357, 363
Russell B. 21, 121, 148, 182, 183, 222, 279, 287–289, 296, 297, 312

Sagan D. 238, 245
Sainsbury R. M. 157, 182, 204, 215
Salmon N. 155, 157, 159, 182
Sarte J. P. 187
Sattig T. 335, 349
Schaffer J. 38, 39, 41, 57, 69, 82, 251, 264
Scheler M. 184, 187, 191, 193–195, 202
Schmidtz D. 465, 468
Schnieder B. 59, 60, 62, 66, 70, 81, 82
Schopenhauer A. 59, 65, 82
Scott D. 476
Sellars W. 57
Shapiro S. 351, 354, 363
Sharma R. K. 394, 396
Sher G. 37, 38, 45, 49, 50, 53, 56, 57, 148
Sider T. 38–41, 52, 53, 57, 122, 127, 149, 164, 182, 251, 255, 264, 308, 312
Simchen O. 119, 123, 135, 139, 140, 149
Simons P. 52, 57, 148, 166, 182, 196, 202, 233, 243–245
Slote M. 460, 465, 468
Snowdon P. F. 265, 275
Sobel J. H. 476
Socrates 221
Sorabji R. 351, 363
Sosa E. 42, 50, 57
Specht E. K. 184, 202
Spinoza B. 389, 391, 392
Stalnaker R. 307, 312
Steglich-Petersen A. 11
Steinberg A. 66, 82
Stojanovic I. 347, 349
Strassmann J. 238, 245
Strawson P. 154, 162, 182
Strawson P. F. 121, 149
Stump E. 31, 388, 396, 413, 415–417, 419, 420, 424, 426, 428, 432, 433, 435–438, 440, 441, 468
Svoboda D. 148
Swenson P. 444, 460, 468
Swoyer C. 285, 288, 296
Szathmáry E. 239, 245
Szatkowski M. IX, 1, 11, 469

Taieb H. 184, 202
Tait W. 352
Tanaka K. 229
Tarski A. 122, 140, 299, 312
Textor M. 200, 202
Thompson N. 42, 52, 57
Tieszen R. 148
Tomasetta A. 265
Troelstra A. S. 363
Trogdon K. 40, 56
Tucker Ch. 461, 462, 465, 468
Turner J. 59, 81, 148, 157, 182
Tye M. 343, 349

Unger P. 329, 349

Vallentyne P. 453, 456, 468
Vallicella W. F. 70, 82, 183, 202
Van Cleve J. 142, 149
Van Fraassen B. 18–21, 307, 312
Van Inwagen P. 11, 60, 66, 67, 70, 71, 74, 76, 79, 82, 85, 87, 89, 91, 93, 99, 101, 103, 105, 108, 110, 113–116, 118–120, 122, 140, 142, 149, 157, 160, 178,

182–184, 202, 209, 215, 233, 242, 245, 247–249, 253, 255, 258, 263–266, 269, 275, 279, 280, 282, 284–287, 296–299, 301, 302, 304, 307, 310, 312, 315–317, 319, 322, 328, 341, 342, 349, 378, 397, 399, 401–403, 405, 407–410, 412, 413, 418, 441, 462, 468
Varzi A. 166, 181
von Solodkoff T. 318, 328

Wójtowicz A. 346, 349
Walton K. 203, 205, 207, 215
Warfield T. A. 81
Weatherson B. 335, 349
Weber Z. 229
Weyl H. 358, 359, 363
Wiggins D. 164, 182, 249, 264, 310, 313
Wilde J. 221, 229
Williams D. C. 164, 182, 313
Williams J. R. G. 332, 349
Williamson T. 333, 349
Wilson J. A. 237, 245
Wippel J. F. 386, 397
Wittgenstein L. 17, 191, 217–219, 221, 222, 227
Wolterstorff N. 105, 118
Worrall J. 251, 264

Yagisawa T. 203, 216
Young J. Z. 248, 264

Zalta E. N. 155, 157, 182, 209, 216
Zimmerman D. 17, 149, 255, 264, 443, 444, 451, 468
Zimmerman M. J. 468

Subject Index

Abstract Artifact Theory 203, 205, 206, 208, 210
actualism 356
actuality 85
adequacy
– material 298
adverbialism 300, 301
almost-identity 332
analysis
– functional 257, 258
animal
– human 266
animalism 265, 266
anty-realism 132
argument
– from contingency 399, 407
– from expert testimony 399
– no-best-world 443
– Rowe's 447
atom
– mereological 20
atomism 255
atonement 430, 433
axiogenesis 377, 385

being 85, 86, 219, 220
– intentional 185
– mental 185
– omnipotent 448
– omniscient 447
belief
– philosophical 409, 410
big contingency 67, 69

category 87
– ontological 114
character 104, 107
character$_+$ 108
choice
– timeless 450
compatibilism 15
composition 233, 247, 248
constitutionalism 103, 105, 107, 111, 117
constitutionalism$_+$ 109, 111, 112, 117
cooperativity 65
correlationism 192, 193, 195, 196
counterpart
– non-relational 286

decomposition
– functional 257
– mechanical 257
degree of being 156
degree of belief 401
dependence 38
determinism 11, 12, 15, 60, 76
– metaphysical 13
dia-philosophy 97
difference
– metaphysical 83
disclosure
– full 403
– partial 403

endurantism 164
engagement
– *de dicto* 213
– *de re* 213
esse
– intelligibile 183
– naturale 183
essence 197
evil 413, 414, 416, 418
exemplification 303
existence 84, 85, 153, 154, 168, 184, 188
– contingent 367, 368, 372
– fictional 11
– necessary 367
explanation 41, 111
– holistic 52

fact
– Moorean 104
– testimonial 404
finegrainedness 65

https://doi.org/10.1515/9783110664812-027

foreknowledge
– simple 450
fundamentality 38

grace
– God's 431, 433
– operative 431, 432
ground 61–63, 65, 106
grounding 38, 42, 45, 112, 123
– holistic 47, 51
– metaphisical 40

holism
– coherentist 48
– foundational 47–49
hylomorphism 247, 250

idealism
– mereological 234
identity 13
– qualitative 322
ideology 319
imagination
– *de re* 213
incompatibilism 404
individual
– emergent 253
infinity
– actual 352, 353
– potential 351–353, 358
instantiation 303

justification 44, 45

kind
– natural 320, 321

language
– natural 155, 160
law
– bivalence 387
locationism 307, 309
logic
– paraconsistent 223, 225

metaphysics 11
– descriptive 154
– fundational 154
– traditional 37
modality 11, 66
modes of being 153, 183–185, 187, 191, 197

naturalism
– ontological 256, 259
necessitarianism 60, 66
necessity 72
nihilism
– compositional 233
nominalism 11, 297
– mereological 306, 307
non-epistemicity 65
nonexistence 177
noophelia 385

o-role 279, 290
object
– abstract 203, 315, 319
– fictional 203
– intentional 177
– material 11
– nonexistent 159
objectuality 183
ontology 120, 122
– constituent 16
– relational 16
order
– relational 287
organism 234, 237

paradox
– König 227, 228
paraphrase 316, 318, 319, 321
part
– functional 249, 255
– temporal 301
person
– human 266
philosophy
– analytic 183
– continental 183
platonism 297
pluralism
– metaphysical 94–96

potentialism 356
– liberal 356
– strict 357
precisification
– admissible 333
predicate
– existence 160–163, 172, 175
predication 209
primitivism 105
primitivism$_+$ 109
Principle
– Pareto 453, 456, 458, 460
principle
– Burley's 387, 388
– explanation 372
– expression 460, 462, 465
– externality of explanation 378, 379
– Hume-Edwards 368, 370
– optimality 374, 375, 384, 394
– Sherlock Holmes 373
Principle of Humility 405
Principle of Sufficient Reason 60
problem
– remnant-person 265, 267, 274
– remnant-thinker 272, 274
problem of the many 329
productivity 65
profile
– predicative 128
property 298, 300
– coinvergence 353
– emergent 254
property role 298, 303, 305
proposition 221

quantifier
– existential 157

realism 11
reality 218
realizability 351, 361
realm
– abundant 290
– sparse 290
reference 299
relationism 105
relationism$_+$ 109

sameness
– qualitative 322
sanctification 435
semantics
– compositional 159
– formal 119, 120
set
– fuzzy 341, 342
– Orłowska's rough 344
– rough 343, 345
sitting 221
Special Composition Question 15, 233, 239, 248
squashing example 251
structure
– ontological 110
substance 372
suffering 418–420, 424, 430
superior
– epistemic 400, 402, 408
supervaluationism 332, 336

testimony 402
– Lewis's 404
theism
– open 450
theory
– of properties 299
thing 372
transference 76
truth
– scrambled 119
truth-making 351

universalism
– compositional 233

whole
– non-living 255
will
– free 11, 12, 15, 60, 65, 76
workshop
– metaphysical 119
world
– maximally good 452